US 20170119861A1

(19) **United States**

(12) **Patent Application Publication** (10) **Pub. No.: US 2017/0119861 A1**

KAKKIS et al. (43) **Pub. Date: May 4, 2017**

(54) **METHODS AND COMPOSITIONS FOR THE TREATMENT OF AMYLOIDOSIS**

(71) Applicant: **Ultragenyx Pharmaceutical Inc.**, Novato, CA (US)

(72) Inventors: **Emil D. KAKKIS**, San Rafael, CA (US); **Michel Claude VELLARD**, San Rafael, CA (US); **Andrzej SWISTOWSKI**, Petaluma, CA (US)

(21) Appl. No.: **15/338,242**

(22) Filed: **Oct. 28, 2016**

Related U.S. Application Data

(60) Provisional application No. 62/248,713, filed on Oct. 30, 2015.

Publication Classification

(51) **Int. Cl.**

A61K 38/48	(2006.01)
C12N 9/48	(2006.01)
C12N 9/64	(2006.01)
A61K 45/06	(2006.01)

(52) **U.S. Cl.**

CPC ***A61K 38/488*** (2013.01); ***A61K 45/06*** (2013.01); ***C12N 9/485*** (2013.01); ***C12N 9/6472*** (2013.01); ***C12N 9/6478*** (2013.01); ***A61K 38/4813*** (2013.01); ***A61K 38/4873*** (2013.01); ***C12Y 304/16001*** (2013.01); ***C12Y 304/22001*** (2013.01); ***C12Y 304/23005*** (2013.01)

(57) **ABSTRACT**

Methods and compositions for the treatment or prevention of amyloidosis are provided. In some embodiments, the methods comprise administering to the subject a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof. Such methods and compositions may be employed to reduce, prevent, degrade and/or eliminate amyloid formation in the lysosome and/or extracellularly.

FIG. 1A

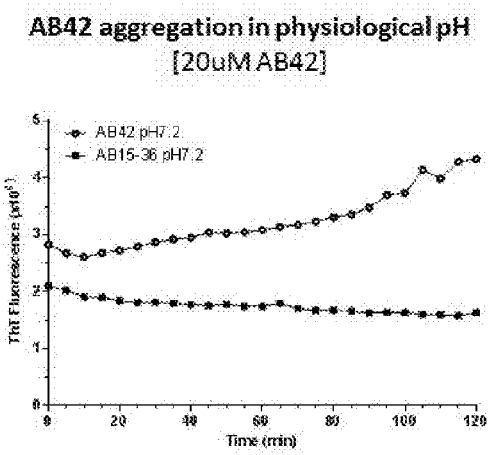


FIG. 1B

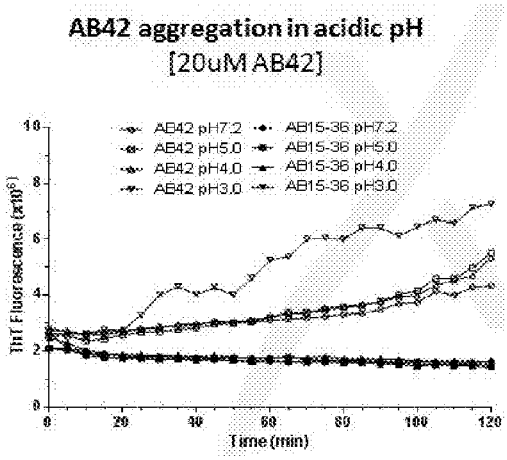


FIG. 1A-B

FIG. 2A

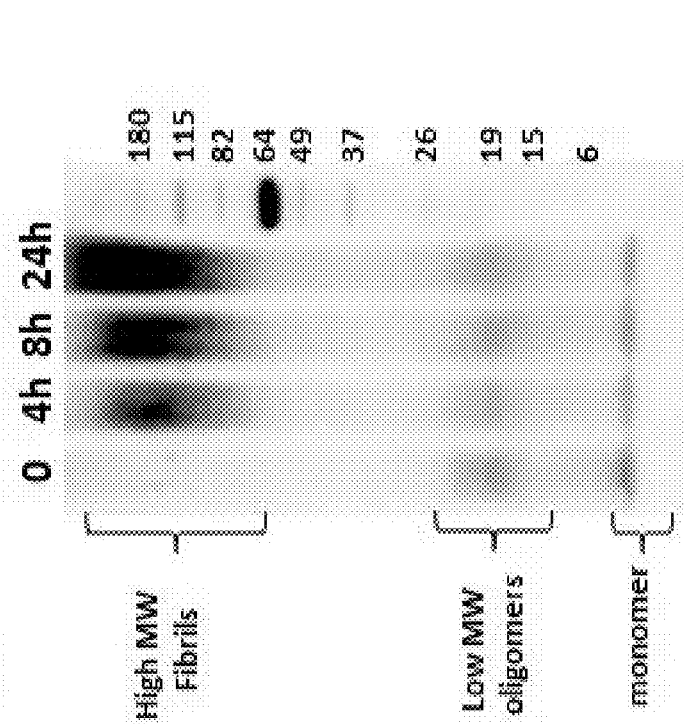


FIG. 2B

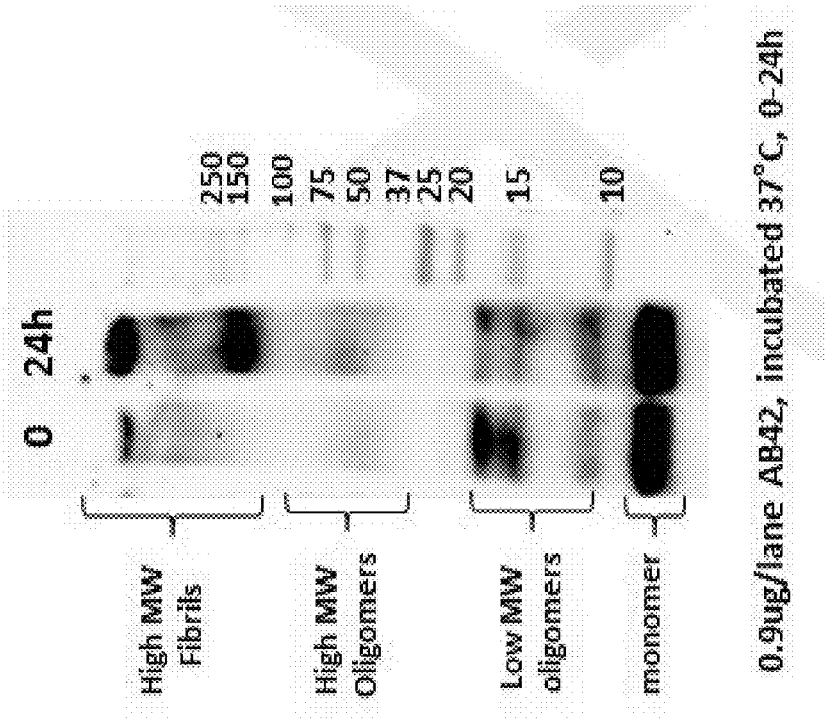


FIG. 2A-B

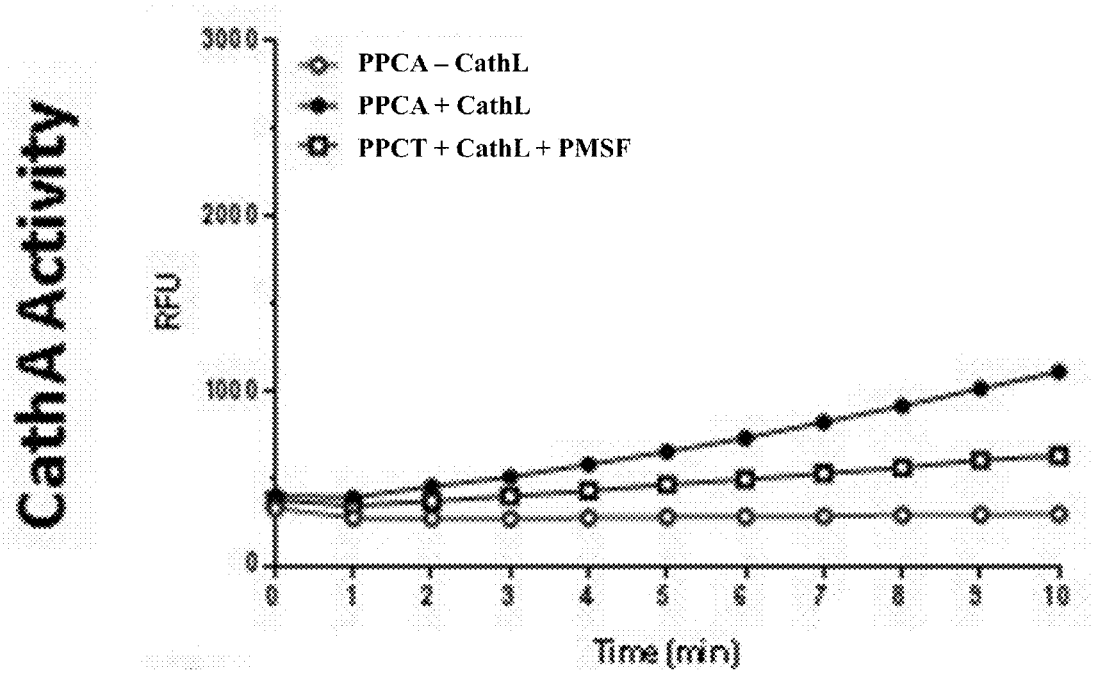

CathA Activity
RFU
300.0
200.0
100.0
0
PPCA – CathL
PPCA + CathL
PPCT + CathL + PMSF
0
1
2
3
4
5
6
7
8
9
10
Time (min)

FIG. 3A

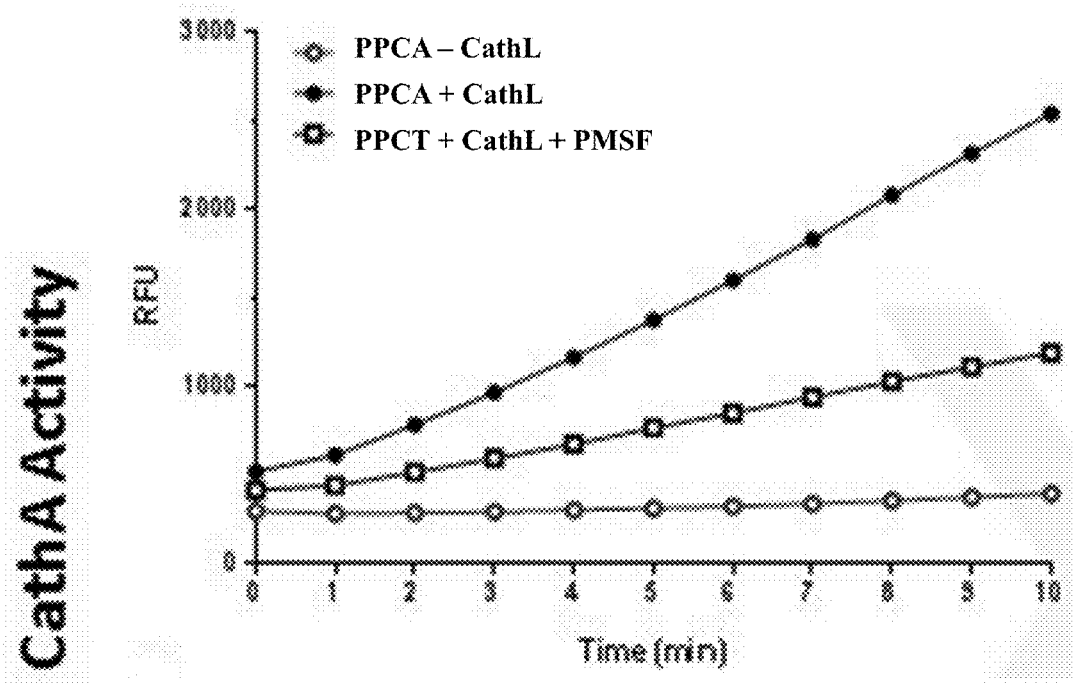

PPCA – CathL
PPCA + CathL
PPCT + CathL + PMSF
CathA Activity
RFU
3000
2000
1000
0
0
1
2
3
4
5
6
7
8
9
10
Time (min)

FIG. 3B

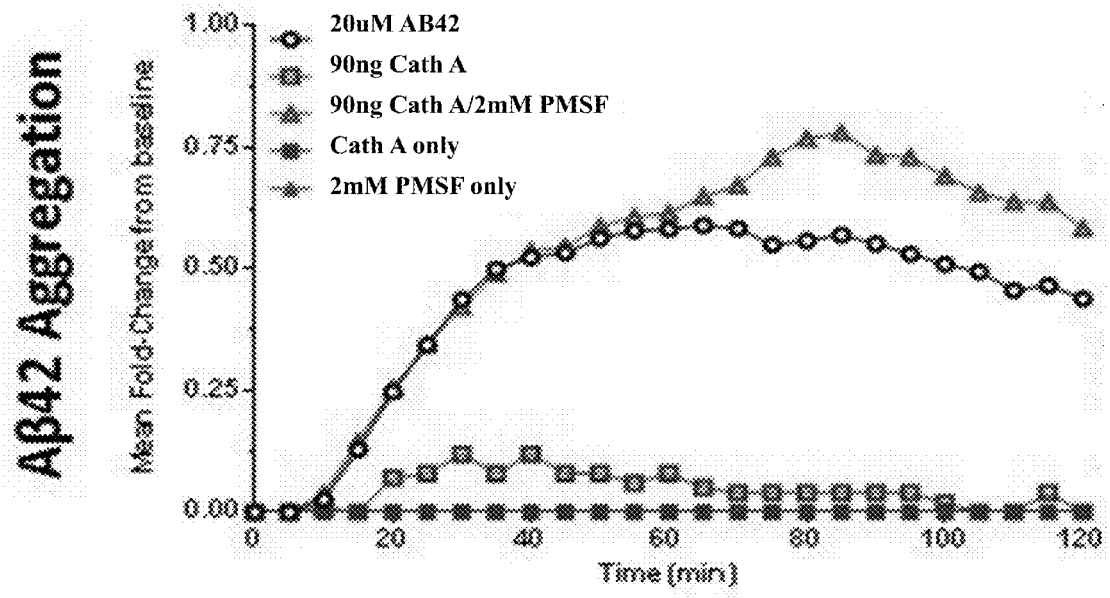

Aβ42 Aggregation
Mean Fold-Change from baseline
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
100
120
Time (min)
20uM AB42
90ng Cath A
90ng Cath A/2mM PMSF
Cath A only
2mM PMSF only

FIG. 3C

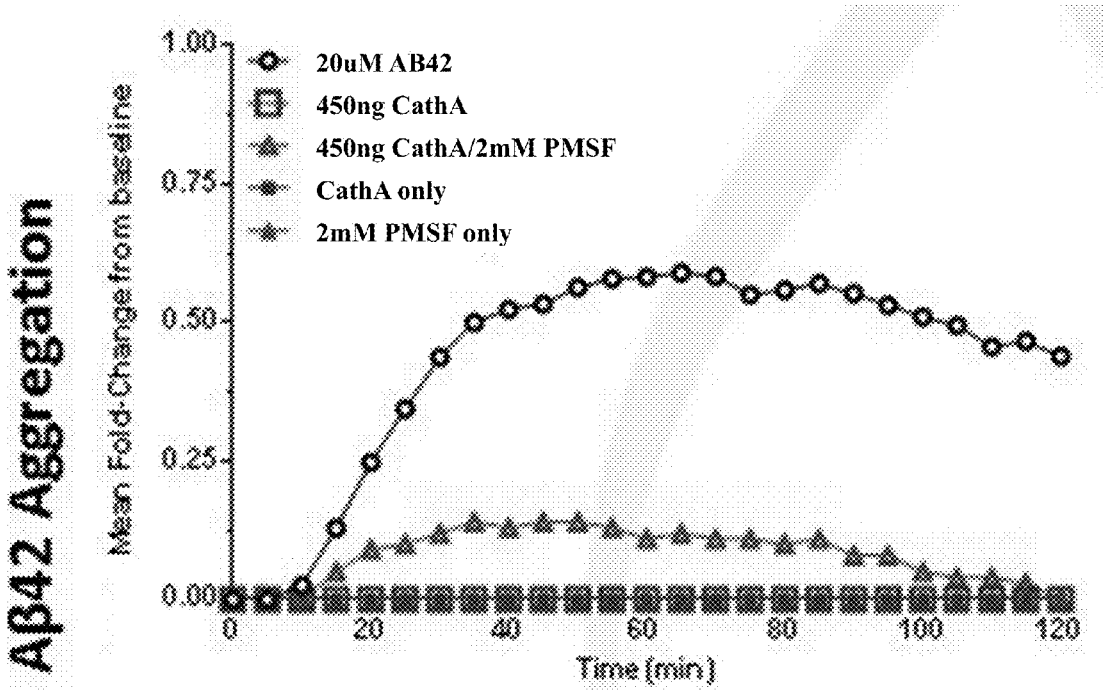

Aβ42 Aggregation
Mean Fold-Change from baseline
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
100
120
Time (min)
20uM AB42
450ng CathA
450ng CathA/2mM PMSF
CathA only
2mM PMSF only

FIG. 3D

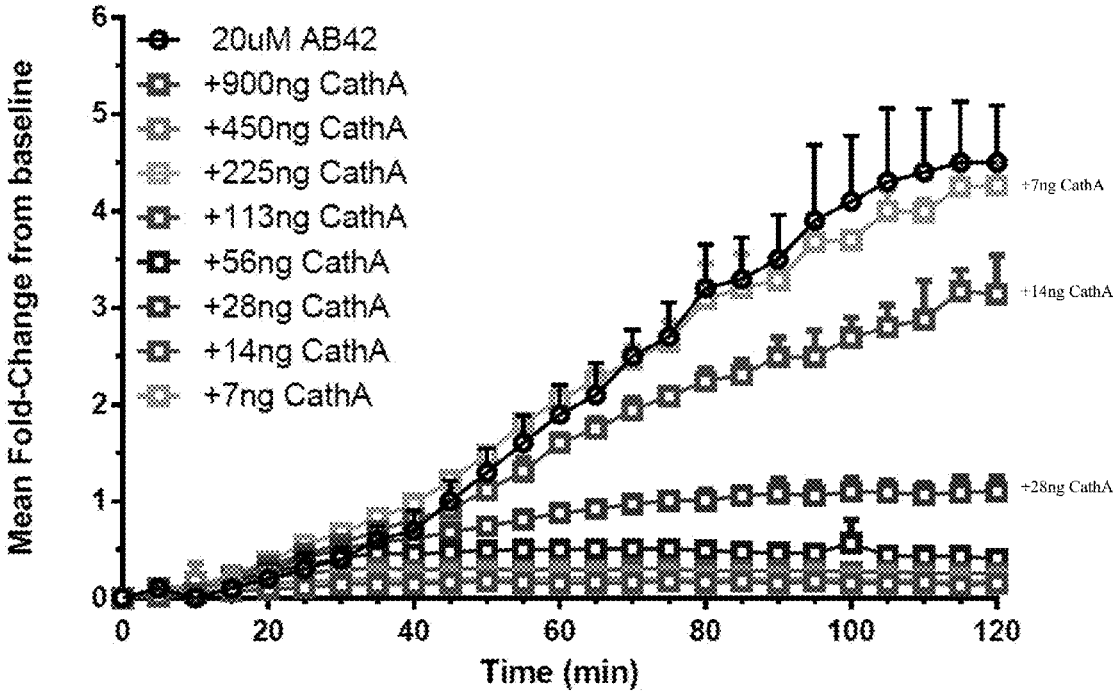

Mean Fold-Change from baseline
Time (min)
20uM AB42
+900ng CathA
+450ng CathA
+225ng CathA
+113ng CathA
+56ng CathA
+28ng CathA
+14ng CathA
+7ng CathA
+7ng CathA
+14ng CathA
+28ng CathA
0
1
2
3
4
5
6
0
20
40
60
80
100
120

FIG. 4A

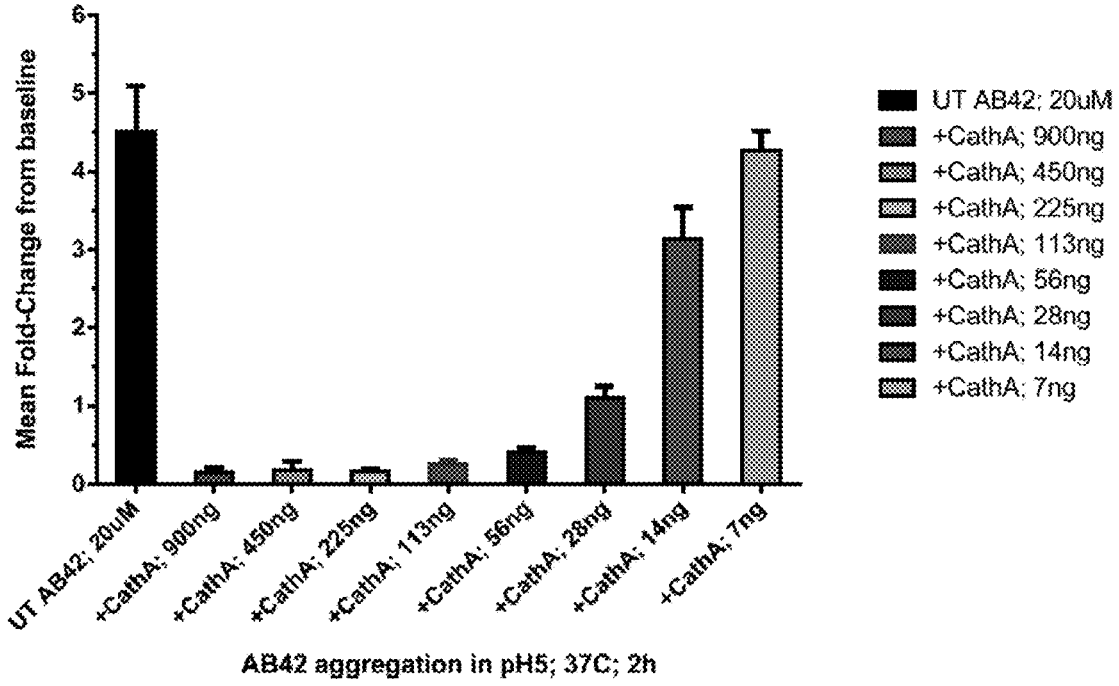

Mean Fold-Change from baseline
0
1
2
3
4
5
6
UT AB42; 20uM
+CathA; 900ng
+CathA; 450ng
+CathA; 225ng
+CathA; 113ng
+CathA; 56ng
+CathA; 28ng
+CathA; 14ng
+CathA; 7ng
AB42 aggregation in pH5; 37C; 2h

FIG. 4B

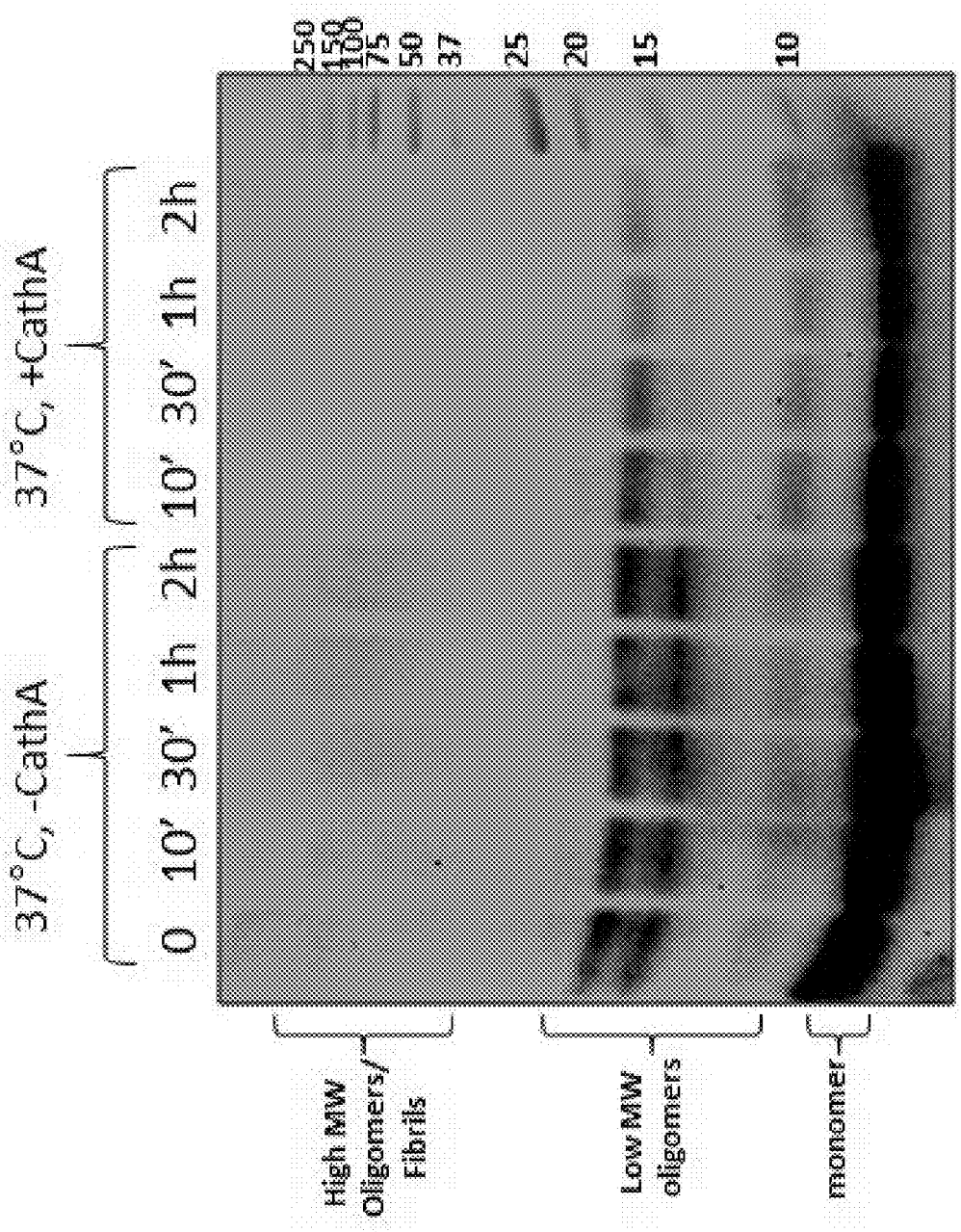

37°C, -CathA
37°C, +CathA
0
10′
30′
1h
2h
10′
30′
1h
2h
250
150
100
75
50
37
25
20
15
10
High MW Oligomers/ Fibrils
Low MW oligomers
monomer

FIG. 5

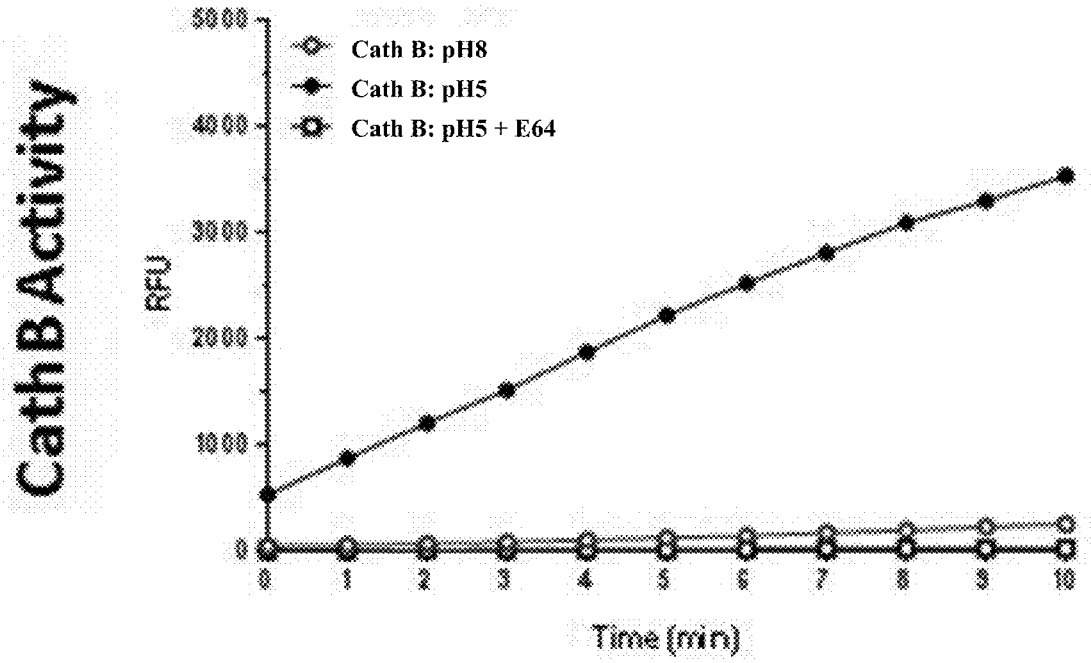

Cath B Activity
RFU
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
7
8
9
10
Time (min)
Cath B: pH8
Cath B: pH5
Cath B: pH5 + E64

FIG. 6A

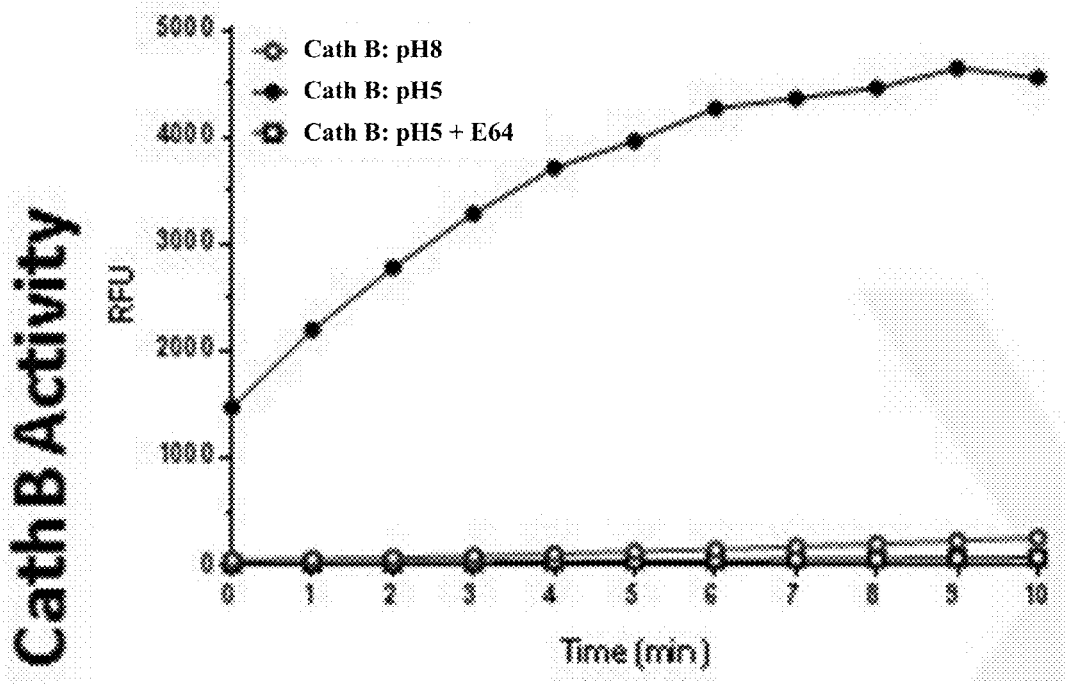

Cath B Activity
RFU
5000
4000
3000
2000
1000
0
0
1
2
3
4
5
6
7
8
9
10
Time (min)
Cath B: pH8
Cath B: pH5
Cath B: pH5 + E64

FIG. 6B

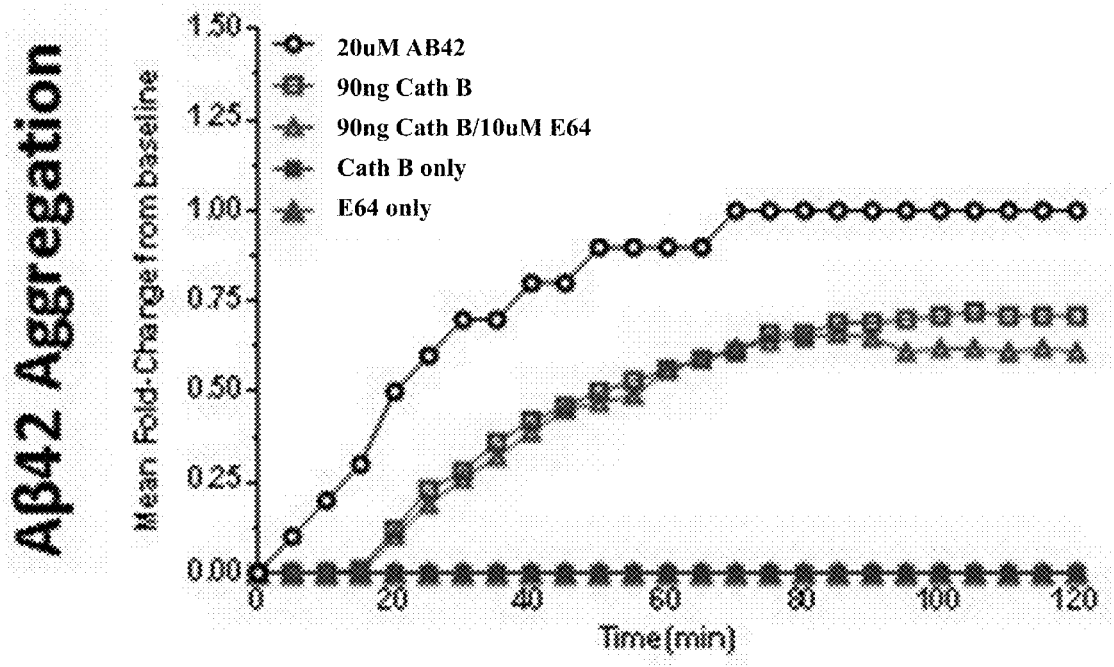

Aβ42 Aggregation
Mean Fold-Change from baseline
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
100
120
Time(min)
20uM AB42
90ng Cath B
90ng Cath B/10uM E64
Cath B only
E64 only

FIG. 6C

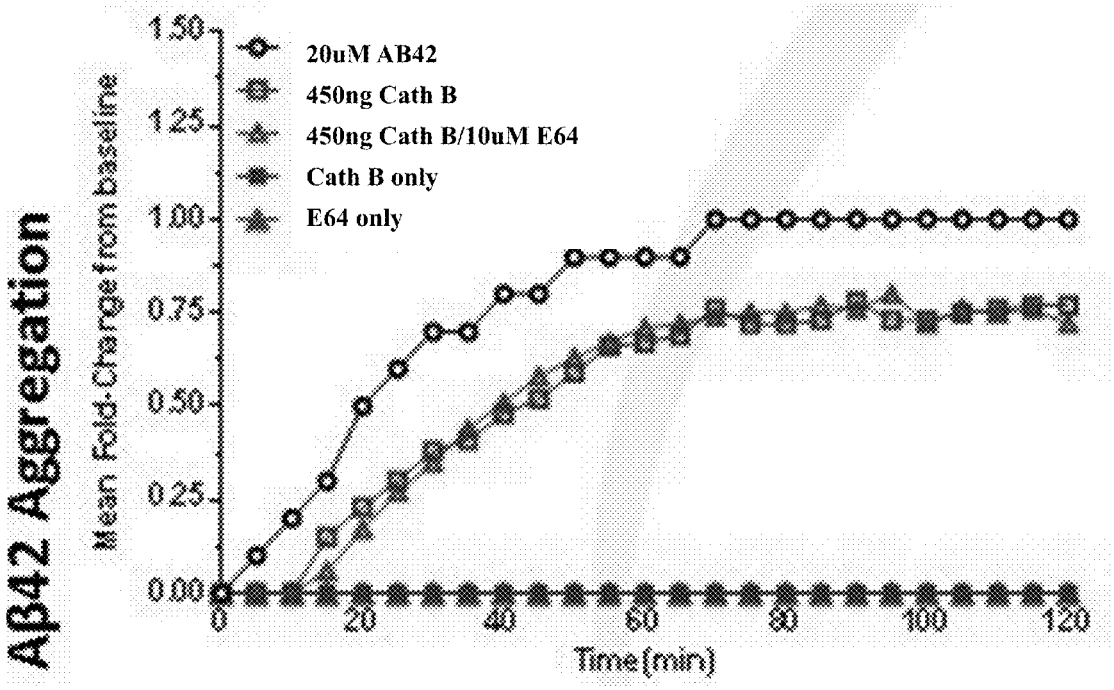

Aβ42 Aggregation
Mean Fold-Change from baseline
1.50
1.25
1.00
0.75
0.50
0.25
0.00
0
20
40
60
80
100
120
Time (min)
20uM Aβ42
450ng Cath B
450ng Cath B/10uM E64
Cath B only
E64 only

FIG. 6D

FIG. 7A

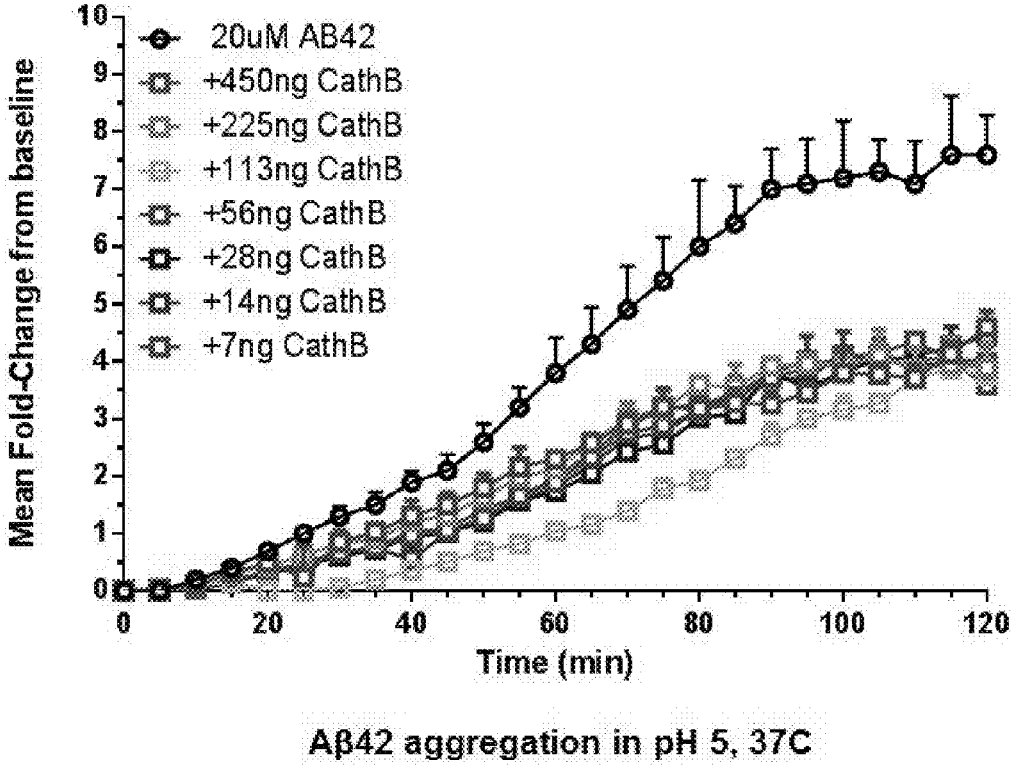

Mean Fold-Change from baseline
Time (min)
20uM AB42
+450ng CathB
+225ng CathB
+113ng CathB
+56ng CathB
+28ng CathB
+14ng CathB
+7ng CathB
0
1
2
3
4
5
6
7
8
9
10
0
20
40
60
80
100
120
Aβ42 aggregation in pH 5, 37C

FIG. 7A

FIG. 7B

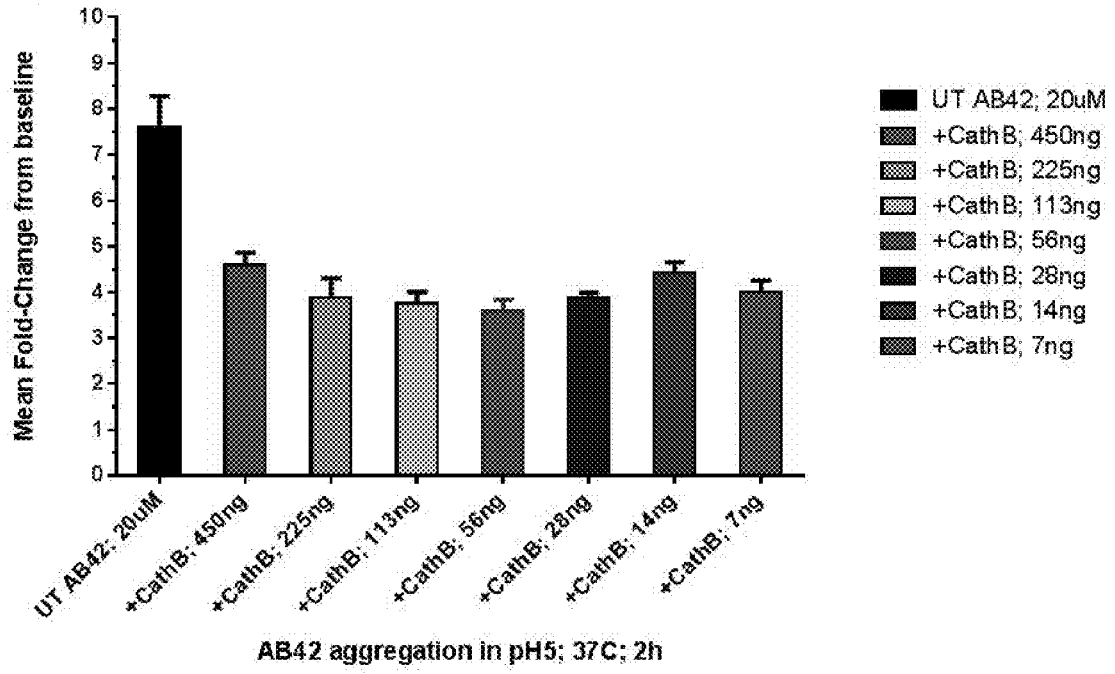

Mean Fold-Change from baseline
10
9
8
7
6
5
4
3
2
1
0
UT AB42; 20uM
+CathB; 450ng
+CathB; 225ng
+CathB; 113ng
+CathB; 56ng
+CathB; 28ng
+CathB; 14ng
+CathB; 7ng
AB42 aggregation in pH5; 37C; 2h

FIG. 7B

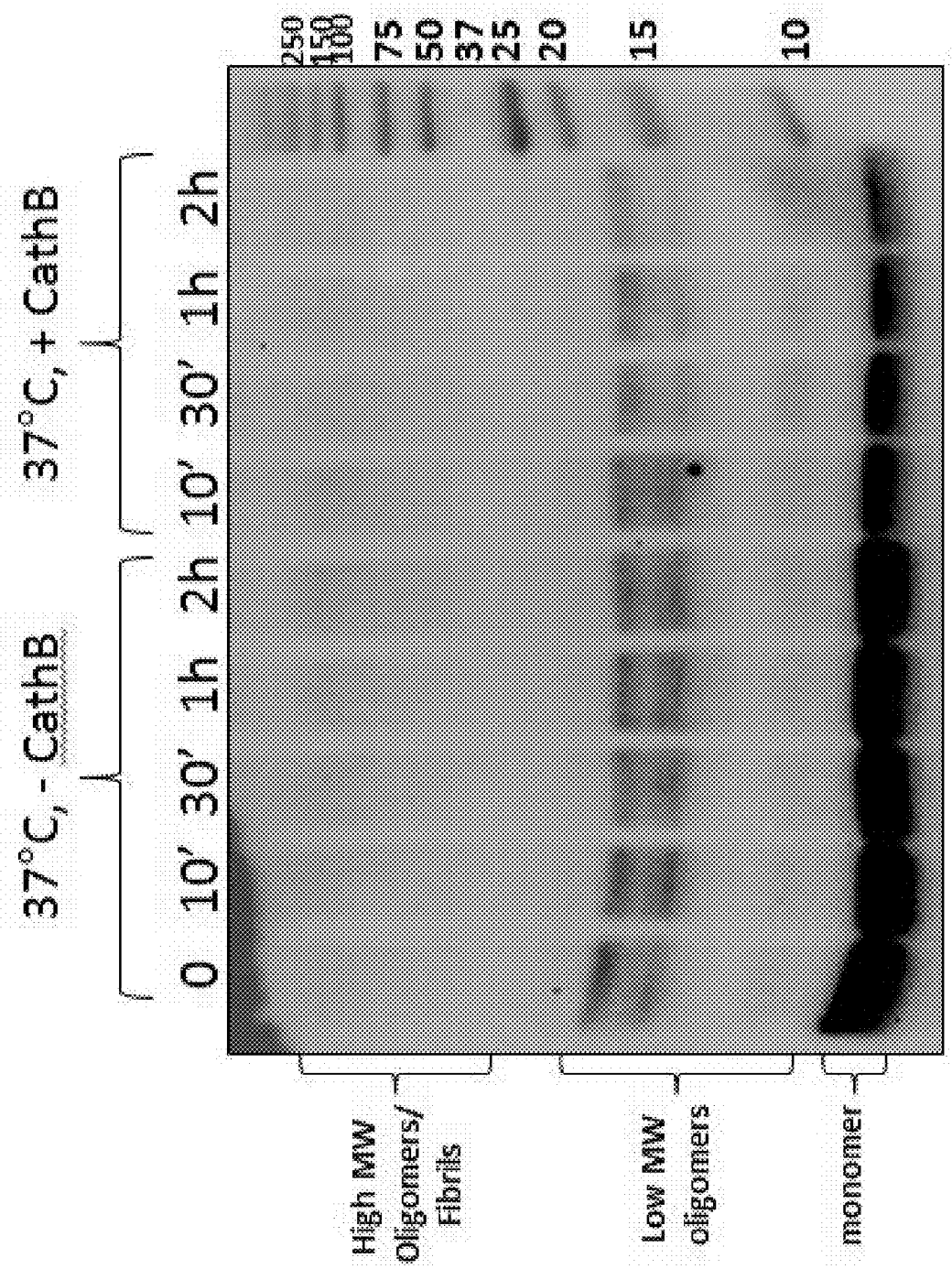

37°C, - CathB
37°C, + CathB
0 10′ 30′ 1h 2h 10′ 30′ 1h 2h
250
150
100
75
50
37
25
20
15
10
High MW Oligomers/ Fibrils
Low MW oligomers
monomer

FIG. 8

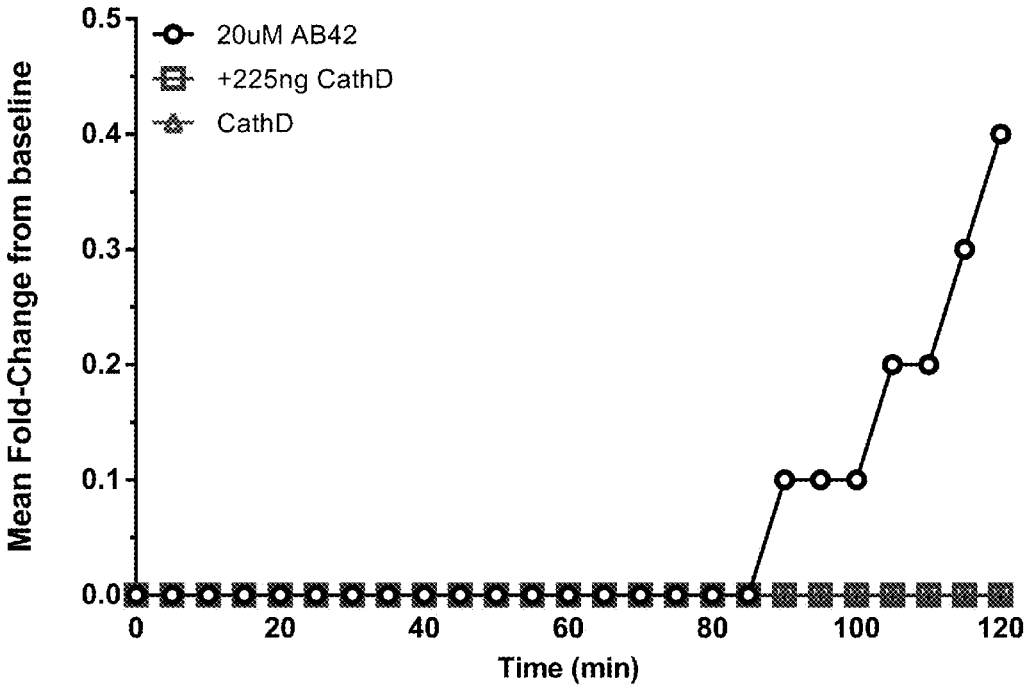

20uM AB42
+225ng CathD
CathD
Mean Fold-Change from baseline
0.5
0.4
0.3
0.2
0.1
0.0
0
20
40
60
80
100
120
Time (min)

FIG. 9

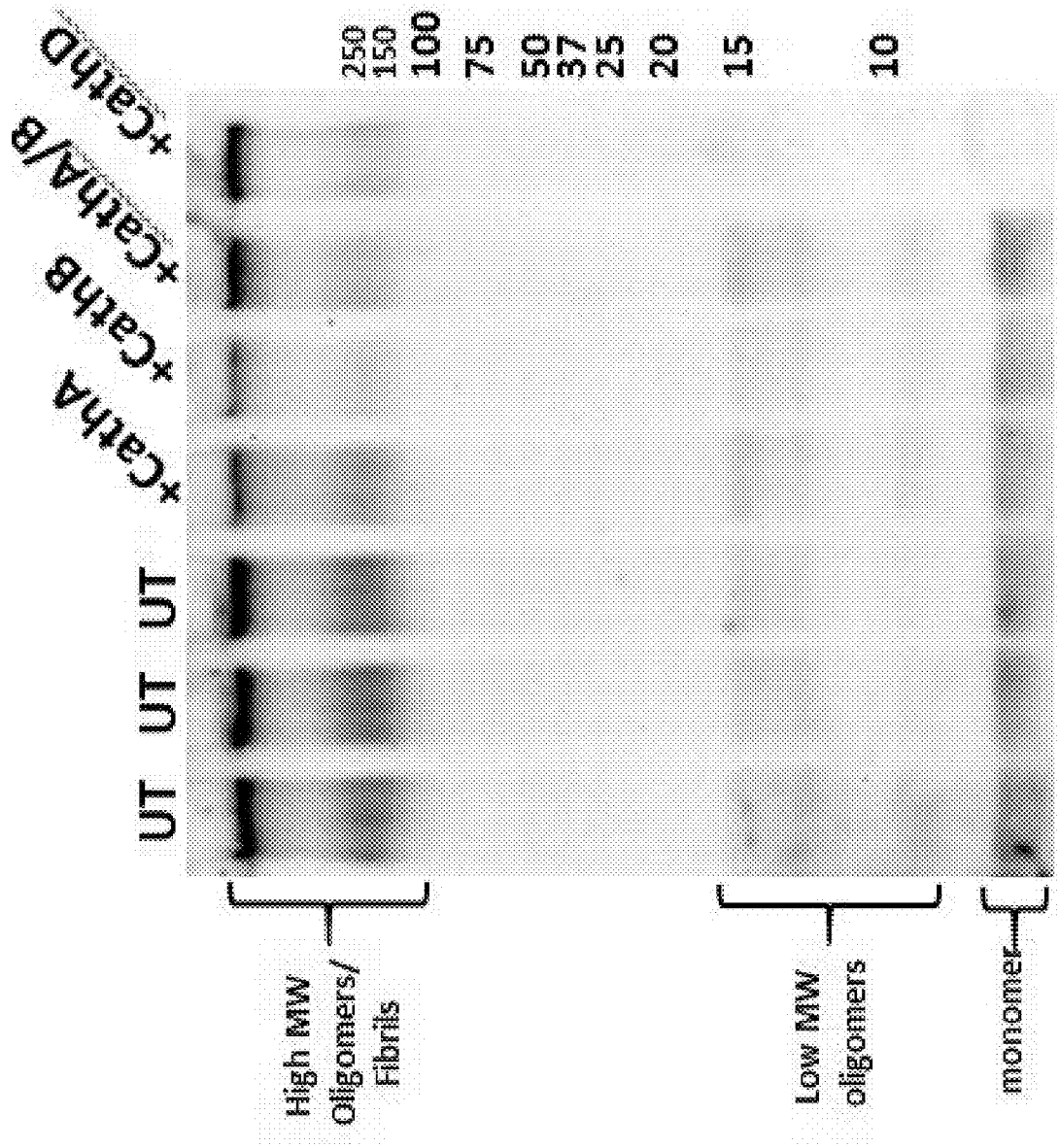

UT
UT
UT
+cathA
+cathB
+cathA/B
+cathD
250
150
100
75
50
37
25
20
15
10
High MW
Oligomers/
Fibrils
Low MW
oligomers
monomer

FIG. 10

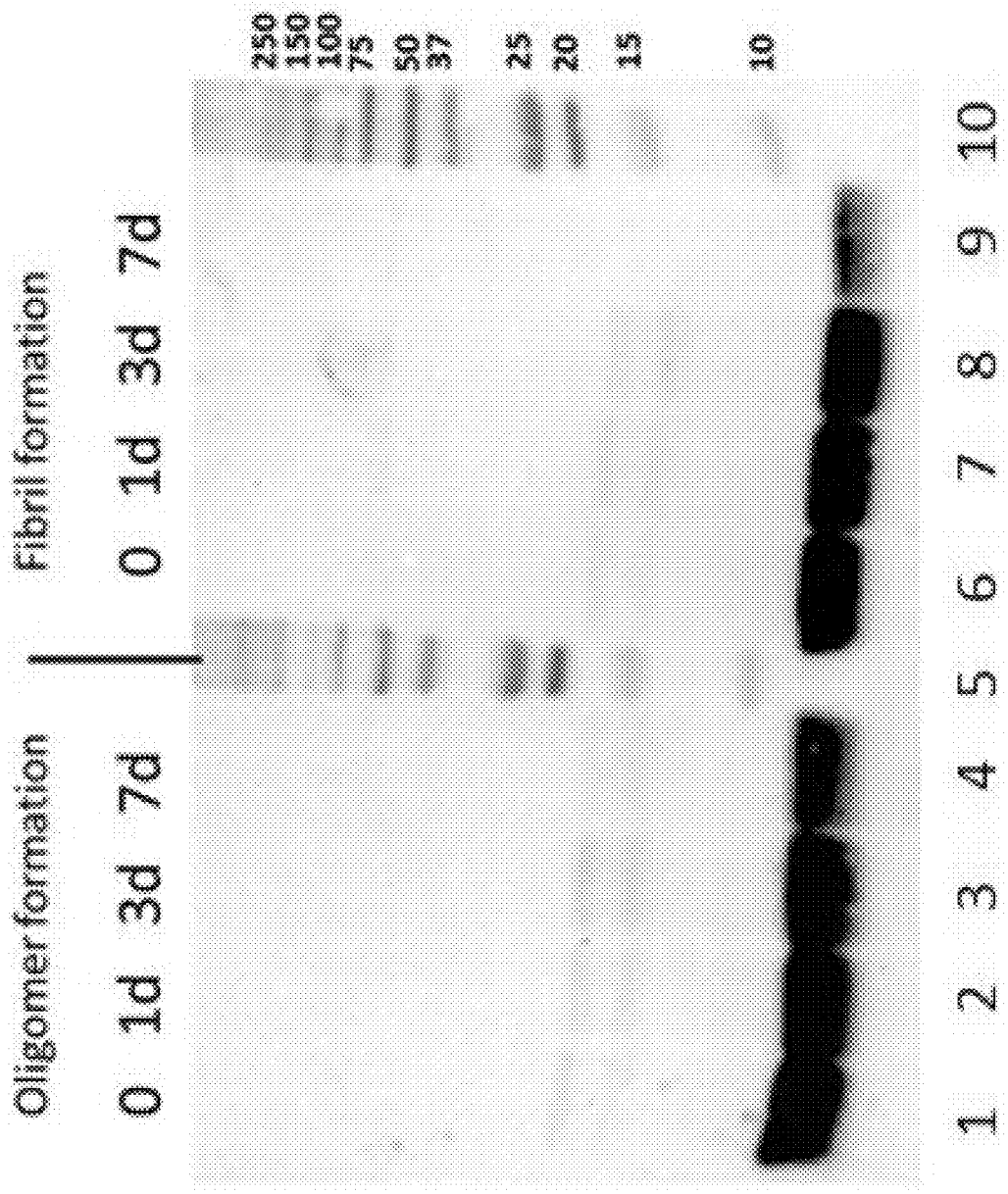

Oligomer formation
0 1d 3d 7d
Fibril formation
0 1d 3d 7d
250
150
100
75
50
37
25
20
15
10
1 2 3 4 5 6 7 8 9 10

FIG. 11

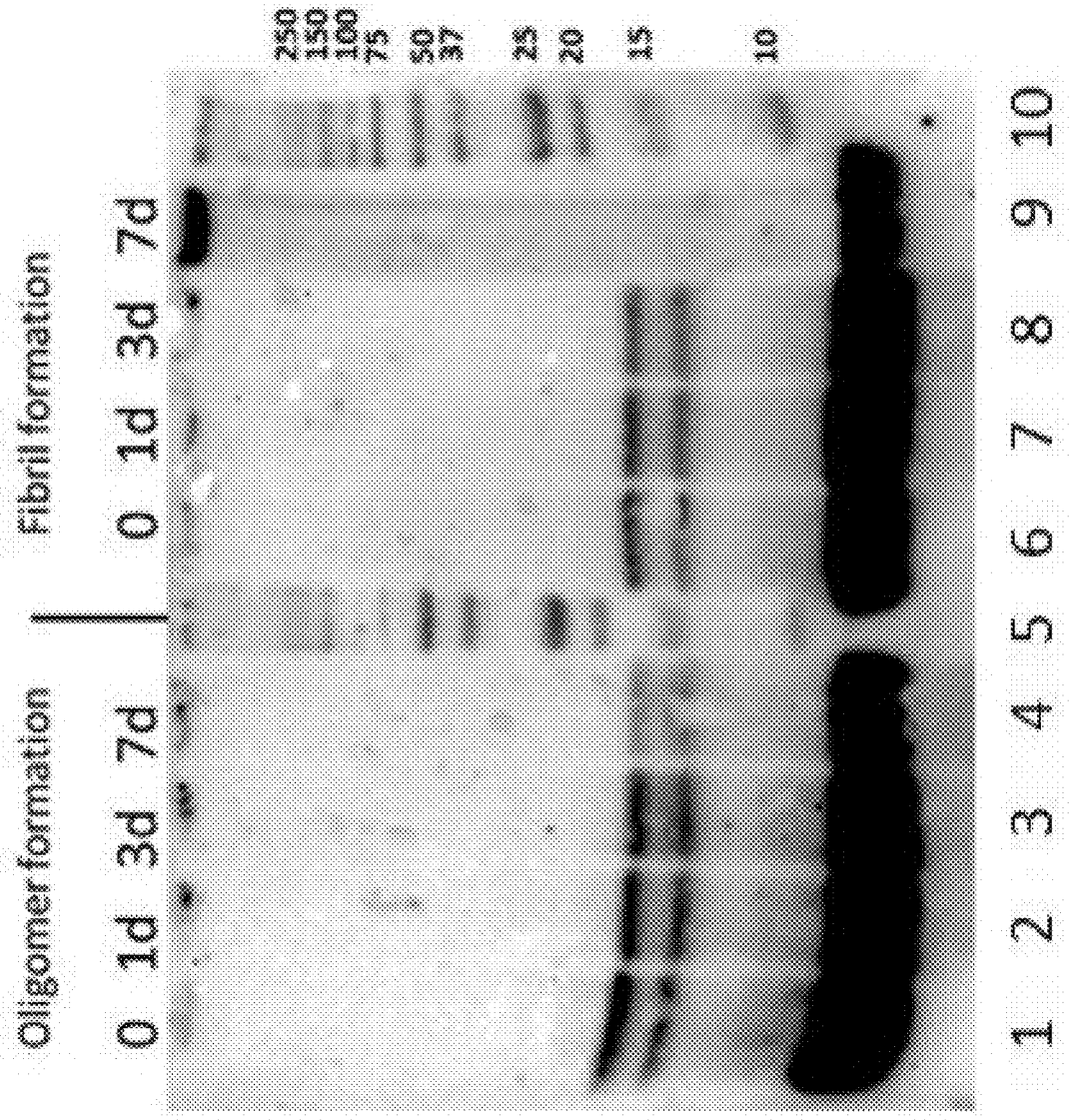
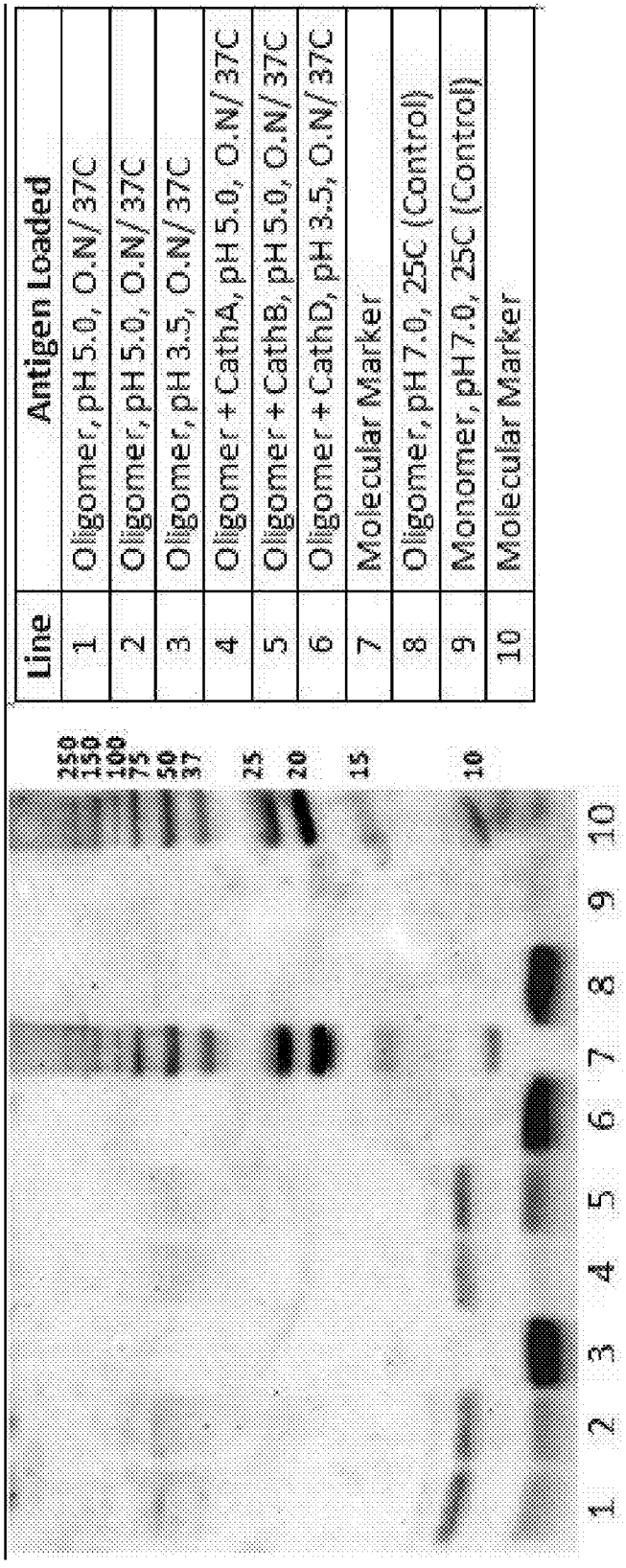

Oligomer formation
Fibril formation
0 1d 3d 7d
0 1d 3d 7d
250
150
100
75
50
37
25
20
15
10
1 2 3 4 5 6 7 8 9 10

FIG. 12



Line	Antigen Loaded
1	Oligomer, pH 5.0, O.N/37C
2	Oligomer, pH 5.0, O.N/37C
3	Oligomer, pH 3.5, O.N/37C
4	Oligomer + CathA, pH 5.0, O.N/37C
5	Oligomer + CathB, pH 5.0, O.N/37C
6	Oligomer + CathD, pH 3.5, O.N/37C
7	Molecular Marker
8	Oligomer, pH 7.0, 25C (Control)
9	Monomer, pH 7.0, 25C (Control)
10	Molecular Marker

FIG. 13

Line	Antigen Loaded
1	Fiber, pH 5.0, O.N/ 37C
2	Fiber, pH 5.0, O.N/ 37C
3	Fiber, pH 3.5, O.N/ 37C
4	Fiber + CathA, pH 5.0, O.N/ 37C
5	Fiber + CathB, pH 5.0, O.N/ 37C
6	Fiber + CathD, pH 3.5, O.N/ 37C
7	Molecular Marker
8	Fiber, pH 7.0, 25C (Control)
9	Monomer, pH 7.0, 25C (Control)
10	Molecular Marker

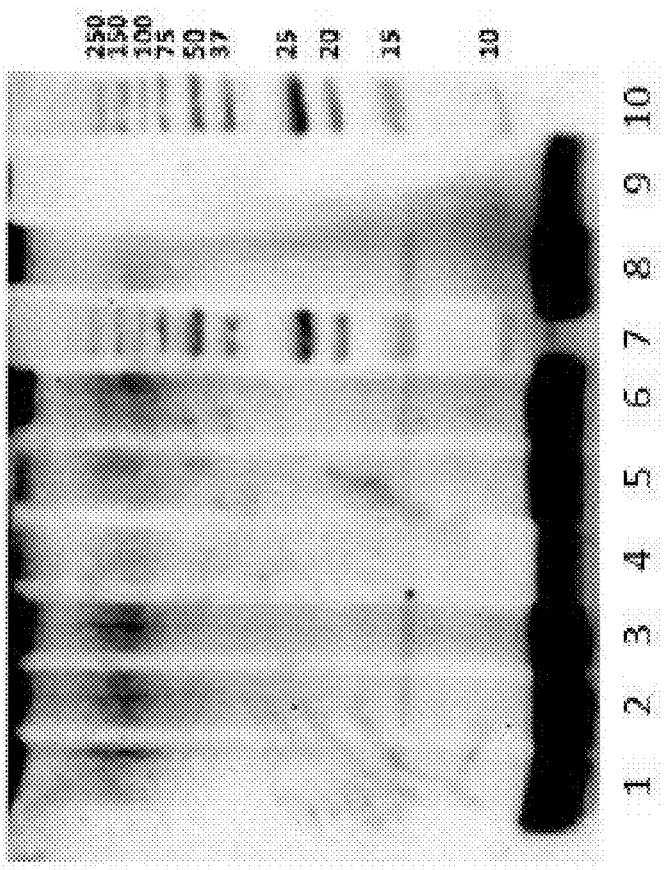


FIG. 14

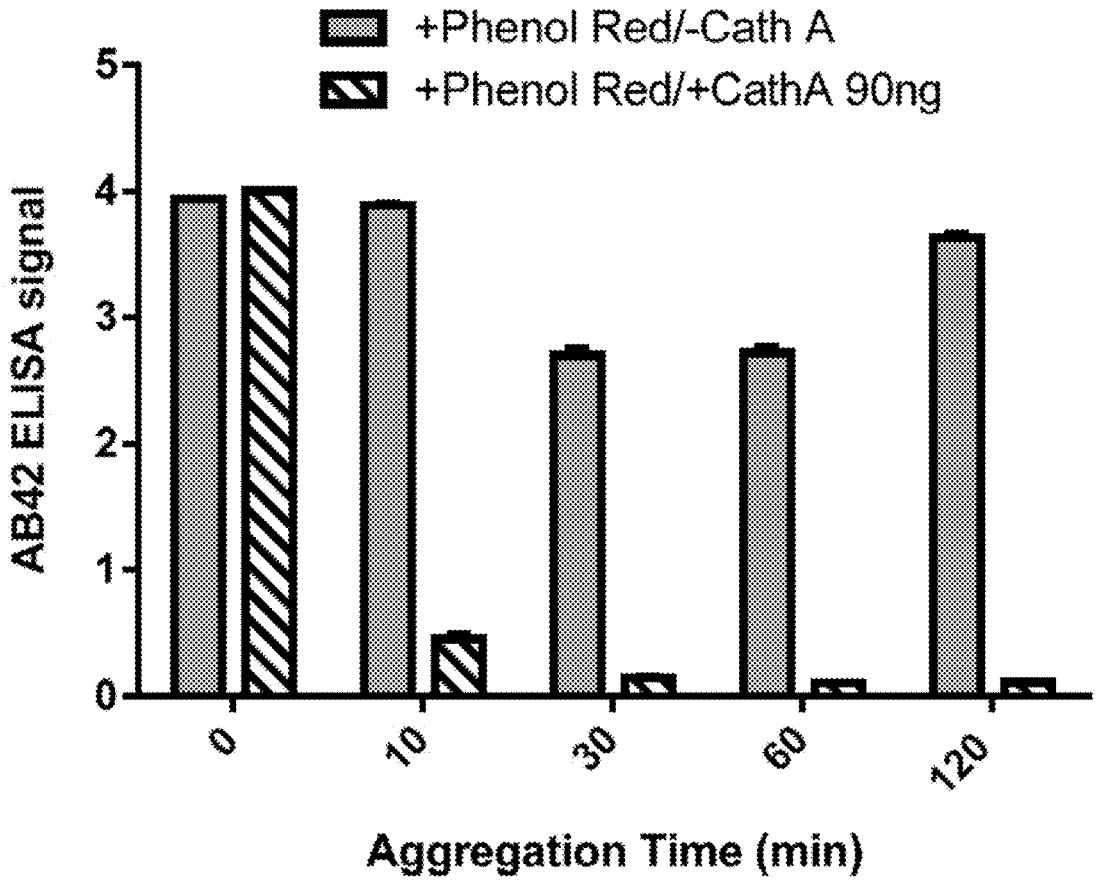

+Phenol Red/-Cath A
+Phenol Red/+CathA 90ng
AB42 ELISA signal
0
1
2
3
4
5
0
10
30
60
120
Aggregation Time (min)

FIG. 15

FIG. 16A

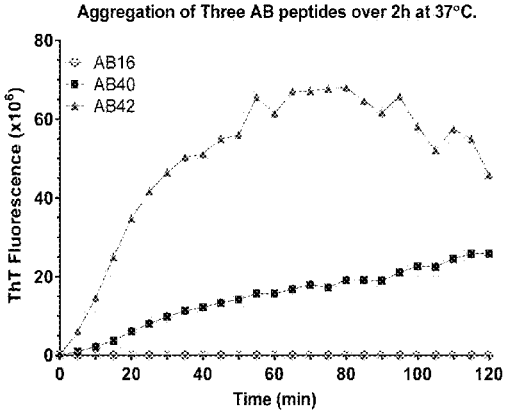


FIG. 16B

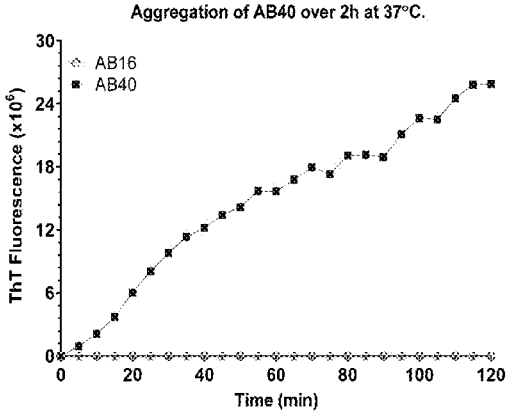


FIG. 16A-B

FIG. 17A

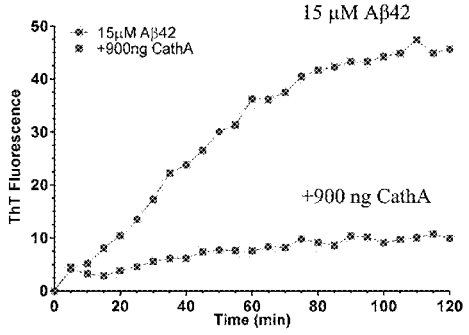


FIG. 17B

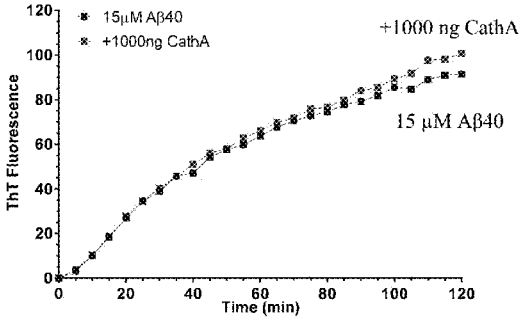


FIG. 17C

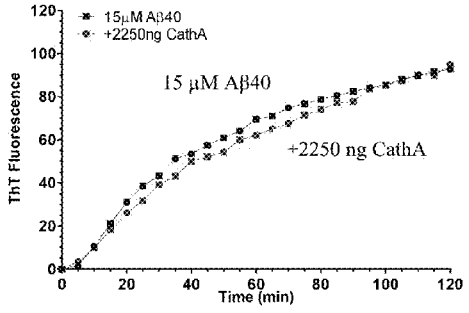


FIG. 17A-C

FIG. 18A

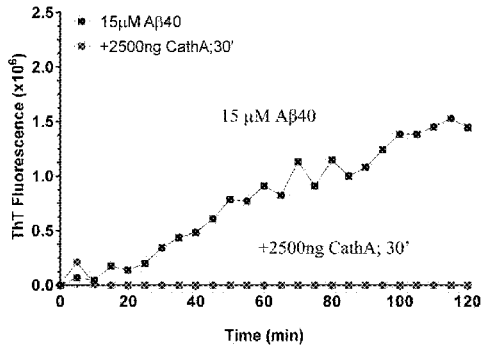


FIG. 18B

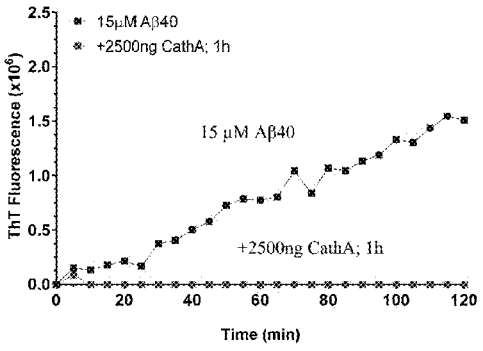


FIG. 18C

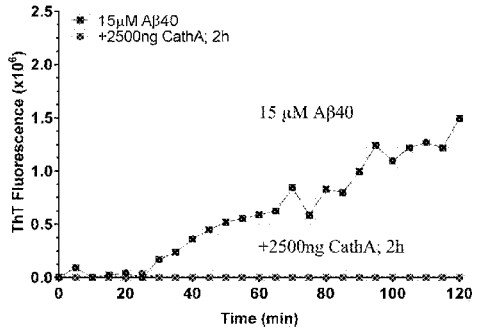


FIG. 18A-C

FIG. 19A

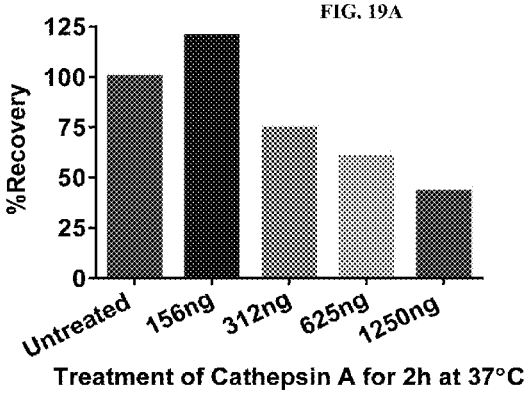


FIG. 19B

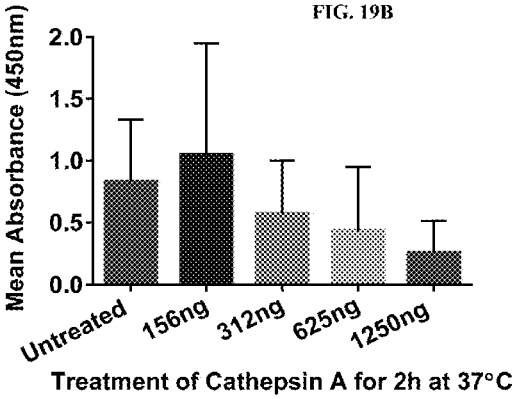


FIG. 19A-B

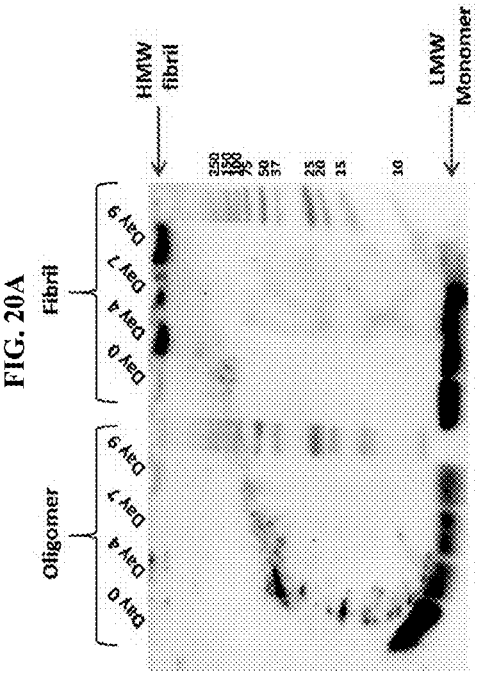


FIG. 20A

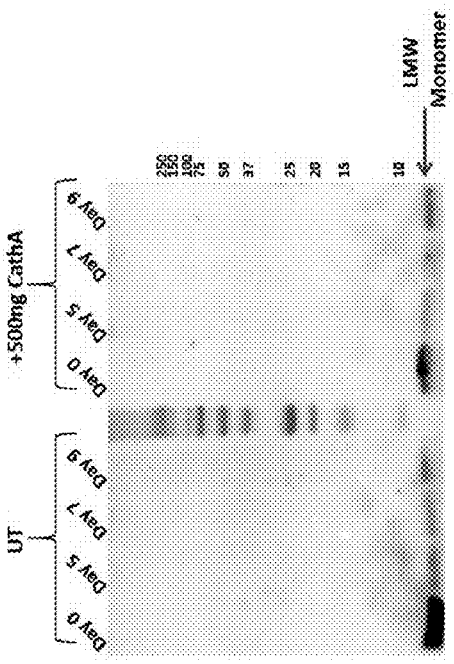


FIG. 20C

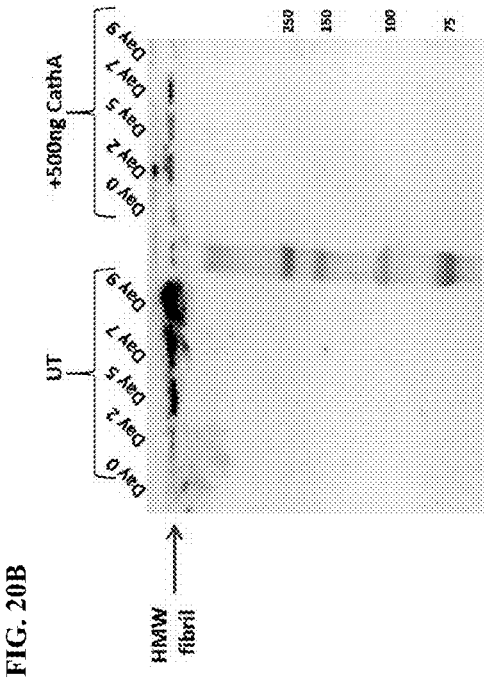


FIG. 20B

FIG. 20A-C

METHODS AND COMPOSITIONS FOR THE TREATMENT OF AMYLOIDOSIS

CROSS-REFERENCE TO RELATED APPLICATIONS

[0001] This application claims priority to U.S. Provisional Application Ser. No. 62/248,713, filed Oct. 30, 2015, which is herein incorporated by reference in its entirety for all purposes.

TECHNICAL FIELD

[0002] The present invention relates to compositions and methods suitable for the prevention or treatment of amyloidosis. For instance, catabolic enzymes are provided to reduce, prevent, or eliminate amyloid formation.

DESCRIPTION OF TEXT FILE SUBMITTED ELECTRONICALLY

[0003] The contents of the text file submitted electronically herewith are incorporated herein by reference in their entirety: A computer readable format copy of the Sequence Listing (filename: ULPI_034_01US_SeqList_ST25.txt, date recorded: Oct. 21, 2016, file size: 146 kilobytes).

BACKGROUND

[0004] Amyloids are insoluble fibrous protein aggregates sharing specific structural traits, e.g., a beta-pleated sheet. They arise from at least 18 inappropriately folded versions of proteins and polypeptides present naturally in the body. These misfolded structures alter their proper configuration such that they erroneously interact with one another or other cell components forming insoluble amyloid fibrils. They have been associated with the pathology of more than 20 serious human diseases. Abnormal accumulation of these amyloid fibrils in organs may lead to amyloidosis, and may play a role in various neurodegenerative disorders, as well as other disorders.

[0005] The formation of these fibrils involves a passage through the lysosome where the acidic environment allows the formation of the protein aggregates. The amyloids are then released from the cell by exocytosis or by cell lysis.

[0006] Trying to eliminate specific fibrils has been the objective of significant research on amyloidosis but without success. Current treatment of amyloidosis involves chemotherapy agents or steroids, such as melphalan and dexamethasone. However, such treatment is not appropriate for all patients and is not effective in many cases due to its specificity. Therefore, there is a great need for alternatives that may safely and effectively prevent or treat diseases associated with amyloidosis.

[0007] The present invention solves the problem of how to prevent and stop the formation of excessive amyloids which have a very deleterious activity in the body. The present invention also solves the problem of specificity, and is applicable to different sources of amyloids and not restricted to a specific disease. The present invention also helps the degradation of already formed fibrils by keeping the lysosome more functional and ready to digest fibrils through endocytosis.

SUMMARY OF THE INVENTION

[0008] The present invention provides methods of treating or preventing amyloidosis in a subject. In some embodiments, the methods comprise administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof.

[0009] In some embodiments, the catabolic enzyme is selected from the group consisting of protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L. In some embodiments, the catabolic enzyme acts to prevent the formation of and/or degrade amyloid within the lysosome, i.e., intralysomally. In other embodiments, the catabolic enzyme acts to prevent the formation of and/or degrade amyloid outside the cell, i.e., extracellularly.

[0010] In some embodiments, the catabolic enzyme comprises a PPCA polypeptide, or a biologically active fragment thereof. In some embodiments, the PPCA polypeptide comprises an amino acid sequence with at least 85% sequence identity to SEQ ID NO: 2, 43, or 45, or a biologically active fragment thereof. In some embodiments, the PPCA polypeptide comprises the amino acid sequence of SEQ ID NO: 2, 43, or 45, or a biologically active fragment thereof.

[0011] In some embodiments, the methods comprise administering a composition comprising a vector, wherein the vector comprises a nucleotide sequence encoding at least one catabolic enzyme of the present invention. In some embodiments, the vector is a viral vector. In some embodiments, the catabolic enzyme is PPCA or a biologically active fragment thereof. In some embodiments, the administration of the PPCA catabolic enzyme comprises administration of a vector encoding a nucleotide sequence having at least 85% identity to SEQ ID NO: 1, 42, or 44. In some embodiments, the nucleotide sequence comprises SEQ ID NO: 1, 42, or 44.

[0012] In some embodiments, the catabolic enzyme comprises a NEU1 polypeptide, or a biologically active fragment thereof. In some embodiments, the NEU1 polypeptide comprises an amino acid sequence with at least 85% sequence identity to SEQ ID NO: 4, or a biologically active fragment thereof. In some embodiments, the NEU1 polypeptide comprises the amino acid sequence of SEQ ID NO: 4, or a biologically active fragment thereof.

[0013] In some embodiments, the administration of the NEU1 catabolic enzyme comprises administration of a vector encoding a nucleotide sequence having at least 85% identity to SEQ ID NO: 3. In some embodiments, the nucleotide sequence comprises SEQ ID NO: 3.

[0014] In some embodiments, the catabolic enzyme comprises a TPP1 polypeptide, or a biologically active fragment thereof. In some embodiments, the TPP1 polypeptide comprises an amino acid sequence with at least 85% sequence identity to SEQ ID NO: 6, or a biologically active fragment thereof. In some embodiments, the TPP1 polypeptide comprises the amino acid sequence of SEQ ID NO: 6, or a biologically active fragment thereof.

[0015] In some embodiments, the administration of the TPP1 catabolic enzyme comprises administration of a vector encoding a nucleotide sequence having at least 85% identity to SEQ ID NO: 5. In some embodiments, the nucleotide sequence comprises SEQ ID NO: 5.

[0016] In some embodiments, at least two catabolic enzymes are administered to the subject. In some embodi-

ments, the at least two catabolic enzymes are selected from protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L.

[0017] In some embodiments, the at least two catabolic enzymes comprise PPCA and NEU1.

[0018] In some embodiments, the catabolic enzyme is targeted to the cell lysosome. In other embodiments, the catabolic enzyme is modified to remain outside the cell, i.e., the enzyme is modified to act extracellularly.

[0019] In some embodiments, the catabolic enzyme prevents the accumulation of and/or degrades amyloid in the cell lysosome. In other embodiments, the catabolic enzyme prevents the accumulation of and/or degrades amyloid outside the cell, i.e., extracellularly.

[0020] In some embodiments, the present invention provides a composition comprising at least two catabolic enzymes, wherein the composition comprises at least one catabolic enzyme that is targeted to the cell lysosome and at least one catabolic enzyme that remains outside the cell. In some embodiments, the catabolic enzymes are selected from protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L. In an exemplary embodiment, the present invention provides a composition comprising at least two catabolic enzymes, wherein the composition comprises a PPCA catabolic enzyme that is targeted to the cell lysosome and a PPCA catabolic enzyme that remains outside the cell.

[0021] In some embodiments, the methods further comprise the administration of one or more additional drugs for treating or preventing amyloidosis. In some embodiments, the one or more additional drugs is/are selected from melphalan, dexamethasone, prednisone, bortezomib, lenalidomide, vincristine, doxorubicin, and cyclophosphamide.

[0022] In some embodiments, the methods further comprise the administration of one or more drugs that acidifies the lysosome. In some embodiments, the drug that acidifies the lysosome is selected from an acidic nanoparticle, a catecholamine, a β-adrenergic receptor agonist, an adenosine receptor agonist, a dopamine receptor agonist, an activator of the cystic fibrosis transmembrane conductance regulator (CFTR), cyclic adenosine monophosphate (cAMP), a cAMP analog, and an inhibitor of glycogen synthase kinase-3 (GSK-3).

[0023] In some embodiments, the methods further comprise the administration of one or more drugs that modulates the lysosome. In an exemplary embodiment, the drug is Z-phenylalanyl-alanyl-diazomethylketone (PADK) or a PADK analog, or a pharmaceutically acceptable salt or ester thereof. In some embodiments, the PADK analog is selected from Z-L-phenylalanyl-D-alanyl-diazomethylketone (PdADK), Z-D-phenylalanyl-L-alanyl-diazomethylketone (dPADK), and Z-D-phenylalanyl-D-alanyl-diazomethylketone (dPdADK).

[0024] In some embodiments, the methods further comprise the administration of one or more drugs that promotes autophagy. In an exemplary embodiment, the drug is selected from an activator of peroxisome proliferator-activated receptor gamma coactivator 1-α (PGC-1α), an inhibitor of Lysine (K)-specific demethylase 1A (LSD1) , an agonist of Peroxisome proliferator-activated receptor (PPAR), an activator of Transcription factor EB (TFEB), an inhibitor of mechanistic target of rapamycin (mTOR), and an inhibitor of glycogen synthase kinase-3 (GSK3).

[0025] In some embodiments, the subject is further treated with stem cell transplantation.

[0026] In some embodiments, the administration is parenteral. In some embodiments, the administration is intramuscular, intraperitoneal, or intravenous.

[0027] In some embodiments, any one of the compositions and drugs provided herein comprise a pharmaceutically acceptable carrier.

[0028] In some embodiments, the subject is a mammal. In some embodiments, the subject is a human.

[0029] In some embodiments, the amyloidosis is light-chain (AL) amyloidosis.

[0030] In some embodiments, the AL amyloidosis involves one or more organs selected from the heart, the kidneys, the nervous system, and the gastrointestinal tract.

[0031] In some embodiments, the amyloidosis is amyloid-beta (Aβ) amyloidosis.

[0032] In some embodiments, the Aβ amyloidosis involves one or more organs selected from the brain, the nervous system, and/or involves various muscles, e.g., muscles of the arms and legs. In some embodiments, the Aβ amyloidosis is associated with Alzheimer's disease. In some embodiments, the Aβ amyloidosis is associated with cerebral amyloid angiopathy. In some embodiments, the Aβ amyloidosis is associated with Lewy body dementia. In some embodiments, the Aβ amyloidosis is associated with inclusion body myositis.

BRIEF DESCRIPTION OF THE DRAWINGS

[0033] FIG. **1**A-B shows the aggregation of synthetic Aβ42 peptide and Aβ15-36 peptide (negative control) monitored by Thioflavin-T (THT). FIG. **1**A. Aggregation at physiological conditions. FIG. **1**B. Aggregation at acidic pH.

[0034] FIG. **2**A-B shows the aggregation of synthetic Aβ42 peptide in vitro over a 24 hour time period as detected by western blot. FIG. **2**A. 12% Bis-Tris gel, reducing conditions, probed with 6E10, a commercially available purified anti-β-amyloid antibody that is reactive to amino acid residues 1-16 of beta amyloid. FIG. **2**B. 18% Tris-Glycine gel, reducing conditions, probed with 6E10.

[0035] FIG. **3**A-D show that cathepsin A (interchangeably referred to herein as Cath A or PPCA) prevents the aggregation of Aβ42 amyloid species. FIG. **3**A. Activation of 90 ng cathepsin A by cathepsin L (full black circles). FIG. **3**B. Activation of 450 ng cathepsin A by cathepsin L. FIG. **3**C. Preventive effect of 90 ng PPCA on Aβ42 aggregation and the inhibition of PPCA by the serine protease inhibitor, PMSF (phenylmethylsulfonyl fluoride) FIG. **3**D Preventive effect of 450 ng PPCA on Aβ42 aggregation. Aβ42 peptides were aggregated alone (open circles), with two concentrations of Cath A (open squares) and with combination of Cath A+inhibitor PMSF (open triangles). Cath A only (full squares) and inhibitor PMSF only (full triangles) were incubated with THT reagent and served as negative controls.

[0036] FIG. **4**A-B shows that Cath A (i.e., PPCA) prevents the aggregation of Aβ42 amyloid species in a dose-dependent manner. FIG. **4**A. Graph showing Aβ42 aggregation over 2 hours at pH 5, 37° C. with varying PPCA concentrations (7 ng to 900 ng) as measured by THT. Aβ42 aggregation was measured alone and with serial dilutions of

PPCA. Lines are labeled for clarity. FIG. **4**B. Bar graph showing end-point (2 hrs) Aβ42 aggregation.

[0037] FIG. **5** shows that Cath A (i.e., PPCA) prevents the aggregation of both high and lower molecular weight species of Aβ42 amyloid. Treatment of 0.9 μg Aβ42 monomer with 500 ng PPCA is shown over a time period of 2 hours on an 18% Tris-Glycine gel, under reducing conditions, probed with 6E10.

[0038] FIG. **6**A-D show that cathepsin B (Cath B) prevents the aggregation of Aβ42 amyloid. FIG. **6**A. Activation of 90 ng cathepsin B and its inhibition by the protease inhibitor E64. FIG. **6**B. Activation of 450 ng cathepsin B and its inhibition by E64. FIG. **6**C. Preventive effect of 90 ng cathepsin B on Aβ42 aggregation and the lack inhibition by E64. FIG. **6**D. Preventive effect of 450 ng cathepsin B on Aβ42 aggregation and the lack inhibition by E64. Aβ42 peptides were aggregated alone (open circles), with two concentrations of Cath B (open squares) and with combination of Cath B+inhibitor E64 (open triangles). Cath B only (full squares) and inhibitor E64 only (full triangles) were incubated with THT reagent and served as negative controls.

[0039] FIG. **7**A-B shows that cathepsin B moderately prevents the aggregation of Aβ42 amyloid species in a dose-dependent manner. FIG. **7**A. Graph showing Aβ42 aggregation over 2 hours at pH 5, 37° C. with varying cathepsin B concentrations (7 ng to 900 ng) as measured by THT. Aβ42 aggregation was measured alone and with serial dilutions of cathepsin B. FIG. **7**B. Bar graph showing end-point (2 hrs) Aβ42 aggregation.

[0040] FIG. **8** shows that cathepsin B prevents the aggregation of both low molecular weight species of Aβ42 amyloid and degrades Aβ42 in a time dependent manner. Treatment of 0.9 μg Aβ42 monomer with 200 ng cathepsin B is shown over a time period of 2 hours on an 18% Tris-Glycine gel, under reducing conditions, probed with 6E10

[0041] FIG. **9** shows that cathepsin D prevents the aggregation of Aβ42 amyloid as monitored by THT. Aβ42 peptides were aggregated alone (empty circles) and with cathepsin D (empty squares) over period of 2 hours. Cathepsin D alone (triangles) was incubated with THT reagent and served as a negative control.

[0042] FIG. **10** shows a western blot demonstrating that PPCA, cathepsin B, PPCA plus cathepsin B, and cathepsin D degrade high molecular weight oligomers/fibrils of Aβ42 amyloid. Cathepsin D degrades low molecular oligomers and completely eliminates Aβ42 monomers.

[0043] FIG. **11** shows a western blot demonstrating a comparison in the detection of Aβ42 oligomers and fibrils using an oligomer specific A11 antibody. Aβ42 peptides were subjected to 7 day aggregation protocols specific for oligomers and fibrils. Reduction of oligomer form in fibril formation (line 9) indicates transition of oligomers into fibril form, which is not detected by oligomer specific A11 antibody.

[0044] FIG. **12** shows a western blot demonstrating a comparison in the detection of Aβ42 oligomers and fibrils using an oligomer and fibril specific E610 antibody. Aβ42 peptides were subjected to 7 day aggregation protocols specific for oligomers and fibrils. Fibril formation was not detected in the oligomer specific protocol at day 7 (line 4). Reduction of oligomer form and appearance of fibril form (smear on line 9) was detected in the fibril formation protocol.

[0045] FIG. **13** shows a western blot illustrating the enzymatic degradation of Aβ42 oligomers as probed by the oligomer specific A11 antibody. Lines 1-6 contain day 9 oligomers aggregated at pH 7.0 at 25° C. and additionally treated overnight at 37° C. in enzyme specific pH. Lines 1-3 are not treated with enzymes. Lines 4-6 represent treatment with 90 ng of cathepsin A, B, and D, respectively. Line 8 contains day 9 oligomers aggregated at pH 7.0 at 25° C. Line 9 contains monomers at pH 7.0. Degradation of oligomers by 90 ng of cathepsin A is shown in line 4. 2 μg of material was loaded on each line.

[0046] FIG. **14** shows a western blot illustrating the enzymatic degradation of Aβ42 fibrils as probed by the oligomer and fibril specific antibody E610. Lines 1-6 contain day 9 fibrils aggregated at pH 7.0 at 25° C. and additionally treated overnight at 37° C. in enzyme specific pH. Lines 1-3 are not treated with enzymes. Lines 4-6 represent treatment with 90 ng of cathepsin A, B, and D, respectively. Line 8 contains day 9 fibers aggregated at pH 7.0 at 25° C. Line 9 contains monomers at pH 7.0. Degradation of fibers and oligomers by 90 ng of cathepsin A is shown in line 4. Degradation of fibers by 90 ng of cathepsin B is shown in line 5. 2 μg of material was loaded on each line.

[0047] FIG. **15** shows a human Aβ42 specific ELISA used to monitor the degradation of Aβ42 monomers with cathepsin A. Treatment of Aβ42 monomers with 90 ng of cathepsin A (striped bars) showed degradation from the C-terminus at various time points (0, 10, 30, 60, 120 min), which is reflected in loss of C-terminal capture by capturing antibody and in effect loss of fluorescent signal. In contrast, Aβ42 monomers not treated with cathepsin A showed lack of C-terminal degradation (solid bars), which is reflected in efficient antibody capture and strong fluorescent signal. An inhibitor of amyloid aggregation, phenol red was used in both cases to prevent peptide aggregation, which could affect capture by the C-terminal antibody in ELISA.

[0048] FIG. **16**A-B show aggregation of Aβ40 and Aβ42 measured by THT assay. Aβ40, Aβ42, and Aβ16 were co-incubated with ThT for 2 h at 37° C. to measure the kinetics of aggregation. Aβ42 aggregates more efficiently and faster than Aβ40. FIG. **16**A. Graphical representation aggregation of Aβ peptides on a single scale. FIG. **16**B. Graphical representation of Aβ40 aggregation on a separate scale.

[0049] FIG. **17**A-C show that simultaneous incubation of Aβ40, Cath A, and THT shows no change in Aβ40 aggregation. Increasing concentrations of Cath A were co-incubated with 15 μM Aβ40 and 2 mM ThT for 2 h at 37° C. to measure how Cath A affected the kinetics of Aβ40 aggregation. FIG. **17**A. 900 ng Cath A was co-incubated with Aβ40 and THT. FIG. **17**B. 1000 ng Cath A was co-incubated with Aβ40 and THT. FIG. **17**C. 2250 ng Cath A was co-incubated with Aβ40 and THT.

[0050] FIG. **18**A-C show that Aβ40 pre-incubated with Cath A leads to loss of its aggregation potential as revealed by lack of THT fluorescence. Aβ40 and 2500 ng Cath A were first incubated for 30', 1 h, and 2 h at 37° C. (FIGS. **18**A, **18**B, and **18**C, respectively). Reactions were then co-incubated with ThT for 2 h at 37° C. to measure how Cath A affected the kinetics of Aβ40 aggregation.

[0051] FIG. **19**A-B show detection of cleavage of Aβ40 C-terminal end using a C-terminal capture antibody. Aβ40 peptide was incubated for 2 h at 37° C. at pH 5 with varying concentrations of Cath A. The reaction was transferred to an

ELISA plate pre-coated with a C-terminal capture antibody and was co-incubated with N-terminal detection antibody overnight at 4° C. Error bars are referring to the standard deviation in the OD values. FIG. **19**A. Recovery rate of undigested Aβ40 in samples treated with increased concentrations of Cath A. FIG. **19**B. Mean absorbance at 450 nm of samples in ELISA wells treated with increased concentrations of Cath A.

[0052] FIG. **20**A-C show aggregation and degradation of Aβ40 amyloid measured by Western Blot. FIG. **20**A. Aggregation into amyloid species. Aβ40 was incubated in either Fibril Buffer or Oligomer buffer at RT for 0-9 days. 2 μg of Aβ40 were loaded per lane on an 18% Tris-Glycine gel and transferred to a PVDF membrane. The blot was probed with an Anti-Aβ40 C-terminal primary antibody (G2-10). Aβ40 incubated with Cath A during fibril formation prevents aggregation. Aβ40 was co-incubated with Cath A in fibril buffer at RT for 0-9 days. To observe high molecular weight bands the gel in FIG. **20**B was run on a 7.5% Tris-glycine gel and to see the low molecular weight bands gel in FIG. **20**C was run on an 18% Tris-glycine gel. 2 μg of Aβ40 were loaded into each lane. Each gel was transferred to a PVDF membrane and probed with an Anti-Aβ40 C-terminal primary antibody (G2-10).

DETAILED DESCRIPTION

[0053] As shown herein, the present inventors have discovered that various catabolic enzymes can be used to prevent the formation of and/or degrade various types of amyloid oligomers and fibrils. Because these oligomers and fibrils can contribute to the development of a variety of amyloid-associated diseases and disorders, the present invention is directed to methods and compositions for the treatment or prevention of amyloidosis in a subject.

[0054] Amyloids are insoluble fibrous protein aggregates sharing specific structural traits. The deposition of normally soluble proteins in this insoluble form can lead to cell death and tissue degeneration. To date, 18 different proteins and polypeptides have been identified in disease-associated amyloid deposits. See Westermark et al. ("Nomenclature of amyloid fibril proteins. Report from the meeting of the International Nomenclature Committee on Amyloidosis, Aug. 8-9, 1998. Part 1." Amyloid. 1999 March; 6(1):63-6), which is incorporated by reference in its entirety. The amyloid fibrils are long, straight, unbranched filaments about 40-120 Å in diameter, which bind to physiological dyes such as Congo red and thioflavine T and are resistant to protease digestion.

[0055] As used herein, amyloidosis refers to a disease that results from accumulation of amyloids. Such diseases to be treated or prevented by the present invention include, but are not limited to, systemic AL amyloidosis, Alzheimer's Disease, Diabetes mellitus type 2, Parkinson's disease, Transmissible spongiform encephalopathy e.g. Bovine spongiform encephalopathy, Fatal Familial Insomnia, Huntington's Disease, Medullary carcinoma of the thyroid, Cardiac arrhythmias, Atherosclerosis, Rheumatoid arthritis, Aortic medial amyloid, Prolactinomas, Familial amyloid polyneuropathy, Hereditary non-neuropathic systemic amyloidosis, Dialysis related amyloidosis, Finnish amyloidosis, Lattice corneal dystrophy, Cerebral amyloid angiopathy, Cerebral amyloid angiopathy (Icelandic type), Sporadic Inclusion Body Myositis, Amyotrophic lateral sclerosis (ALS), Prion-related or Spongiform encephalopathies, such as Creutzfeld-Jacob, Dementia with Lewy bodies, Frontotemporal dementia with Parkinsonism, Spinocerebellar ataxias, Spinocerebellar ataxia, Spinal and bulbar muscular atrophy, Hereditary dentatorubral-pallidoluysian atrophy, Familial British dementia, Familial Danish dementia, Non-neuropathic localized diseases, such as in Type II diabetes mellitus, Medullary carcinoma of the thyroid, Atrial amyloidosis, Hereditary cerebral haemorrhage with amyloidosis, Pituitary prolactinoma, Injection-localized amyloidosis, Aortic medial amyloidosis, Hereditary lattice corneal dystrophy, Corneal amyloidosis associated with trichiasis, Cataract, Calcifying epithelial odontogenic tumors, Pulmonary alveolar proteinosis, Inclusion-body myositis, Cutaneous lichen amyloidosis, and Non-neuropathic systemic amyloidosis, such as AL amyloidosis, AA amyloidosis, Familial Mediterranean fever, Senile systemic amyloidosis, Familial amyloidotic polyneuropathy, Hemodialysis-related amyloidosis, ApoAI amyloidosis, ApoAII amyloidosis, ApoAIV amyloidosis, Finnish hereditary amyloidosis, Lysozyme amyloidosis, Fibrinogen amyloidosis, Icelandic hereditary cerebral amyloid angiopathy, familial amyloidosis, and systemic amyloidosis which occurs in multiple tissues, such as light-chain amyloidosis, and other various neurodegenerative disorders. In exemplary embodiments, the amyloidosis is light-chain (AL) amyloidosis. In further exemplary embodiments, the AL amyloidosis involves one or more organs selected from the heart, the kidneys, the nervous system, and the gastrointestinal tract.

[0056] In some embodiments, the present invention provides methods and compositions for the treatment or prevention of a disease associated with amyloidosis in a subject, wherein the disease is associated with the formation of amyloid-beta (Aβ or Abeta) peptides. These peptides result from the amyloid precursor protein (APP), which is cleaved by beta secretase and gamma secretase to yield amyloid-beta. In some embodiments, the disease associated with the formation of amyloid-beta is selected from Alzheimer's Disease, cerebral amyloid angiopathy, Lewy body dementia, and inclusion body myositis.

[0057] In alternative embodiments, the present invention provides methods and compositions for the treatment or prevention of a disease associated with amyloidosis in a subject, wherein the disease is not associated with the formation of amyloid beta, i.e., wherein the disease is a disease other than one associated with the formation of amyloid beta, e.g., a disease other than Alzheimer's disease, cerebral amyloid angiopathy, Lewy body dementia, and inclusion body myositis.

[0058] In one embodiment, the disease associated with amyloidosis is light-chain (AL) amyloidosis. In another embodiment, the disease associated with amyloidosis is selected from Parkinson's Disease, Huntington's Disease, Rheumatoid arthritis, and a prion-related disease.

[0059] In some embodiments, the amyloidosis is a systemic amyloidosis. Systemic amyloidosis encompasses a complex group of diseases caused by tissue deposition of misfolded proteins that result in progressive organ damage.

[0060] As noted above, in some embodiments, the amyloidosis is light-chain (AL) amyloidosis (also known as, i.e. a.k.a., primary systemic amyloidosis (PSA) or primary amyloidosis). AL amyloidosis refers to a condition caused when a subject's antibody-producing cells do not function properly and produce abnormal protein fibers made of components of antibodies called light chains. In some embodi-

ments, such light chains form amyloid deposits in one or more different organs which may cause or already caused damage to these organs. In some embodiments, the abnormal light chains are in blood and/or urine. In some embodiments, the abnormal light chains are "Bence Jones proteins". In some embodiments, the AL amyloidosis affects the heart, peripheral nervous system, gastrointestinal tract, blood, lungs and/or skin. Clinical features of AL amyloidosis also may include a constellation of symptoms and organ dysfunction that can include cardiac, renal, and hepatic dysfunction, gastrointestinal involvement, neuropathies and macroglossia.

[0061] In some embodiments, the amyloidosis is AA amyloidosis (a.k.a. secondary amyloidosis, AA), caused by deposited proteins called serum amyloid A protein (SAA). In some embodiments, the SAA protein is mainly deposited in the liver, spleen and/or kidney. In some embodiments, the AA amyloidosis leads to nephrotic syndrome. In some embodiments, the AA amyloidosis is caused by autoimmune diseases (e.g., Rheumatoid arthritis, Ankylosing spondylitis, or Crohn's disease and ulcerative colitis), Chronic infections (e.g., Tuberculosis, Bronchiectasis, or Chronic osteomyelitis), autoinflammatory diseases (e.g., Familial Mediterranean fever (FMF), Muckle-Wells syndrome (MWS), Cancer (e.g., Hodgkin's lymphoma, Renal cell carcinoma), and/or Chronic foreign body reaction (e.g., Silicone-induced granulomatous reaction).

[0062] In some embodiments, the amyloidosis is familial amyloidosis. In some embodiments, the familial amyloidosis is ATTR amyloidosis (a.k.a. or senile systemic amyloidosis) which is due one or more inherited amyloidosis, such as a mutation in the transthyretin (TTR) gene that produces abnormal transthyretin protein. In some embodiments, the familial amyloidosis is caused by one or more mutation in apolipoprotein A-I (AApoAI), apolipoprotein A-II (AApoAII), gelsolin (AGel), fibrinogen (AFib), lysozyme (ALys), and/or Lect2.

[0063] In some embodiments, the amyloidosis is Beta-2 Microglobulin Amyloidosis (Abeta2m). Beta-2 microglobulin amyloidosis is caused by chronic renal failure and often occurs in patients who are on dialysis for many years. Amyloid deposits are made of the beta-2 microglobulin protein that accumulated in tissues, particularly around joints, when it cannot be excreted by the kidney because of renal failure.

[0064] In some embodiments, the amyloidosis is Localized Amyloidosis (ALoc). In some embodiments, localized amyloid deposits in the airway (trachea or bronchus), eye, or urinary bladder. In some embodiments, the ALoc is caused by local production of immunoglobulin light chains not originating in the bone marrow. In some embodiments, the ALoc is associated with endocrine proteins, or proteins produced in the skin, heart, and other sites. These usually do not become systemic.

[0065] In some embodiments, the amyloidosis occurs in the kidney of the subject. In some embodiments, the amyloidosis in the kidney is AA amyloidosis. In some embodiments, the AA amyloidosis leads to nephrotic syndrome. In some embodiments, the amyloidosis in the kidney is AL amyloidosis. In some embodiments, symptoms of kidney disease and renal failure associated with AL amyloidosis include, but are not limited to, fluid retention, swelling, and shortness of breath.

[0066] In some embodiments, the amyloidosis occurs in the heart of the subject. In some embodiments, the amyloidosis in the heart is AL amyloidosis. In some embodiments, the amyloidosis in the heart leads to heart failure and/or irregular heart beat.

[0067] In some embodiments, the amyloidosis occurs in the gastrointestinal tract of the subject. In some embodiments, symptoms of GI amyloidosis include, but are not limited to, esophageal reflux, constipation, nausea, abdominal pain, diarrhea, weight loss, and early satiety. In some embodiments, the amyloidosis occurs in the duodenum, stomach, colo-rectum, and/or esophagus.

[0068] In some embodiments, the treatment methods provided herein alleviate, reduce the severity of, or reduce the occurrence of, one or more of the symptoms associated with amyloidosis. Such symptoms include those symptoms associated with light-chain (AL) amyloidosis (primary systemic amyloidosis) and/or AA amyloidosis (secondary amyloidosis). In some embodiments, the symptoms include, but are not limited to, fluid retention, swelling, shortness of breath, fatigue, irregular heartbeat, numbness of hands and feet, rash, shortness of breath, swallowing difficulties, swollen arms or legs, esophageal reflux, constipation, nausea, abdominal pain, diarrhea, early satiety, stroke, gastrointestinal disorders, enlarged liver, diminished spleen function, diminished function of the adrenal and other endocrine glands, skin color change or growths, lung problems, bleeding and bruising problems, fatigue and weight loss, decreased urine output, diarrhea, hoarseness or changing voice, joint pain, and weakness. In some embodiments, the symptoms are those associated with amyloid-beta (Aβ) amyloidosis. In some embodiments, the symptoms include, but are not limited to, common symptoms of Alzheimer's disease, including memory loss, confusion, trouble understanding visual images and spatial relationships, and problems speaking or writing.

[0069] According to the methods of the present invention, the term "subject," includes any subject that has, is suspected of having, or is at risk for having a disease or condition. Suitable subjects (or patients) include mammals, such as laboratory animals (e.g., mouse, rat, rabbit, guinea pig), farm animals, and domestic animals or pets (e.g., cat, dog). Non-human primates and human patients are also included. A subject "at risk" may or may not have detectable disease, and may or may not have displayed detectable disease prior to the prevention or treatment methods described herein. "At risk" denotes that a subject has one or more so-called risk factors, which are measurable parameters that correlate with development of any one of the diseases, disorders, conditions, or symptoms described herein,. A subject having one or more of these risk factors has a higher probability of developing any one of the diseases, disorders, conditions, or symptoms described herein than a subject without these risk factor(s). In some embodiments, the subject is a mammal. In some embodiments, the subject is a human. In some embodiments, the subject is a human diagnosed as having amyloidosis or disease/symptom caused by or associated with amyloidosis. In some embodiments, the subject is a human suspected to have amyloidosis. In some embodiments, the subject is a human having high risk of developing amyloidosis. In some embodiments, the subject is an amyloidosis patient with one or more diseases/conditions/symptoms as described herein.

[0070] The terms "treating" and "treatment" as used herein refer to an approach for obtaining beneficial or desired results including clinical results, and may include even minimal changes or improvements in one or more measurable markers of the disease or condition being treated. A treatment is usually effective to reduce at least one symptom of a condition, disease, disorder, injury or damage. Exemplary markers of clinical improvement will be apparent to persons skilled in the art. Examples include, but are not limited to, one or more of the following: decreasing the severity and/or frequency one or more symptoms resulting from the disease, diminishing the extent of the disease, stabilizing the disease (e.g., preventing or delaying the worsening of the disease), delay or slowing the progression of the disease, ameliorating the disease state, decreasing the dose of one or more other medications required to treat the disease, and/or increasing the quality of life, etc.

[0071] "Prophylaxis," "prophylactic treatment," "prevention," or "preventive treatment" refers to preventing or reducing the occurrence or severity of one or more symptoms and/or their underlying cause, for example, prevention of a disease or condition in a subject susceptible to developing a disease or condition (e.g., at a higher risk, as a result of genetic predisposition, environmental factors, predisposing diseases or disorders, or the like).

[0072] The present invention provides methods of treating or preventing amyloidosis in a subject. In some embodiments, the methods comprise administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof. In some embodiments, the methods comprise increasing the expression, activity, and/or concentration of at least one catabolic enzyme in the subject. Increasing the expression, activity, and/or concentration of a given catabolic enzyme may be accomplished at the genomic DNA level, transcriptional level, post-transcriptional level, translational level, and/or post-translational level, including but not limited to, increasing the gene copy number, mRNA transcription rate, mRNA abundance, mRNA stability, protein translation rate, protein stability, protein modification, protein activity, protein complex activity, etc. Increasing the concentration of a given catabolic enzyme may further be accomplished by administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof. As used herein, the term catabolic enzyme refers not only to the natural form the enzyme, but also any purified, isolated, synthetic, recombinant, and functional variants, fragments, chimeras, and mutants of the natural enzyme.

[0073] In some embodiments, the at least one catabolic enzyme is selected from the non-limiting group consisting of protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L.

[0074] In some embodiments, the at least one catabolic enzyme is PPCA (a.k.a. Protective Protein Cathepsin A, PPGB, Carboxypeptidase C, EC 3.4.16.5, GSL, GLB2, Carboxypeptidase Y-Like Kininase, NGBE, carboxypeptidase-L, Protective Protein For Beta-Galactosidase (Galactosialidosis), deamidase, Beta-Galactosidase, Lysosomal Carboxypeptidase A, Beta-Galactosidase Protective Protein, Lysosomal Protective Protein, Protective Protein For Beta-Galactosidase, Urinary Kininase, EC 3.4.168, or Carboxypeptidase L) is classified both as a cathepsin and a carboxypeptidase.

[0075] In some embodiments, the at least one catabolic enzyme is PPCA. PPCA is a glycoprotein that associates with the lysosomal enzymes beta-galactosidase and neuraminidase to form a complex of high-molecular-weight multimers. The formation of this complex provides a protective role for stability and activity. It is protective for β-galactosidase and neuraminidase. In some embodiments, the PPCA can be a natural, synthetic, or recombinant protein. In some embodiments, the PPCA polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 2, 43, or 45. In some embodiments, the PPCA polypeptide comprises the amino acid sequence of SEQ ID NO: 2, 43, or 45.

[0076] In some embodiments, the at least one catabolic enzyme is Neuraminidase 1 (NEU1, a.k.a. sialidase 1, lysosomal sialidase, EC 3.2.1.18, Acetylneuraminyl Hydrolase, SIAL1, Lysosomal Sialidase, exo-alpha-sialidase, NANH, sialidase-1, or G9 Sialidase) is a lysosomal neuraminidase enzyme. NEU1 is an enzyme that cleaves terminal sialic acid residues from substrates such as glycoproteins and glycolipids. In some embodiments, the NEU1 can be a natural, synthetic, or recombinant protein. In some embodiments, the NEU1 polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 4. In some embodiments, the NEU1 polypeptide comprises the amino acid sequence of SEQ ID NO: 4.

[0077] In some embodiments, the at least one catabolic enzyme is Tripeptidyl peptidase 1 (TPP1, Spinocerebellar Ataxia, Autosomal Recessive 7, CLN2, SCAR7, Growth-Inhibiting Protein 1, Cell Growth-Inhibiting Gene 1 Protein, Lysosomal Pepstatin Insensitive Protease, Tripeptidyl Aminopeptidase, Tripeptidyl-Peptidase 1, LPIC, Lysosomal Pepstatin-Insensitive Protease, or EC 3.4.14.9). TPP1 is an enzyme that cleaves N-terminal tripeptides from substrates and has weaker endopeptidase activity. In some embodiments, the TPP1 can be a natural, synthetic, or recombinant protein. In some embodiments, the TPP1 polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 6. In some embodiments, the TPP1 polypeptide comprises the amino acid sequence of SEQ ID NO: 6.

[0078] In some embodiments, the at least one catabolic enzyme is Cathepsin B (a.k.a. EC 3.4.22.1, CPSB, Amyloid Precursor Protein Secretase, Cysteine Protease, APPS, APP secretase, or EC 3.4.22). Cathepsin B is a lysosomal cysteine protease composed of a dimer of disulfide-linked heavy and light chains, both produced from a single protein precursor. In some embodiments, the Cathepsin B can be a natural, synthetic, or recombinant protein. In some embodiments, the Cathepsin B polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%,

87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 8, 47, 49, 51, 53, 55, or 57. In some embodiments, the Cathepsin B polypeptide comprises the amino acid sequence of SEQ ID NO: 8, 47, 49, 51, 53, 55, or 57.

[0079] In some embodiments, the at least one catabolic enzyme is Cathepsin D (a.k.a. EC 3.4.23.5, CTSD). Cathepsin D refers is a lysosomal acid protease active in intracellular protein breakdown. In some embodiments, the Cathepsin D can be a natural, synthetic, or recombinant protein. In some embodiments, the Cathepsin D polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 68. In some embodiments, the Cathepsin D polypeptide comprises the amino acid sequence of SEQ ID NO: 68. In some embodiments, the Cathepsin D polypeptide harbors one or more modifications relative to the amino acid sequence of SEQ ID NO: 68. In certain embodiments, the Cathepsin D polypeptide comprises the amino acid sequence of SEQ ID NO: 68, wherein the polypeptide harbors a modification at an amino acid position selected from position 58 (A to V), position 229 (F to I), position 282 (G to R), and position 383 (W to C).

[0080] In some embodiments, the at least one catabolic enzyme is Cathepsin E (a.k.a. EC 3.4.23.34, CTSE). Cathepsin E is a lysosomal aspartyl protease. In some embodiments, the Cathepsin E can be a natural, synthetic, or recombinant protein. In some embodiments, the Cathepsin E polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 69, 70, or 71. In some embodiments, the Cathepsin E polypeptide comprises the amino acid sequence of SEQ ID NO: 69, 70, or 71. In some embodiments, the Cathepsin E polypeptide harbors one or more modifications relative to the amino acid sequence of SEQ ID NO: 69, 70, or 71. In certain embodiments, the Cathepsin E polypeptide comprises the amino acid sequence of SEQ ID NO: 69, wherein the polypeptide harbors a modification at an amino acid position selected from position 82 (I to V) and position 329 (T to I).

[0081] In some embodiments, the at least one catabolic enzyme is Cathepsin K (a.k.a. EC 3.4.22.38, CTSO, Pycnodysostosis, PYCD, Cathepsis O, PKND, Cathepsin X). Cathepsin K is a lysosomal cysteine protease involved in bone remodeling and resorption, defined by its high specificity for kinins. In some embodiments, the Cathepsin K can be a natural, synthetic, or recombinant protein. In some embodiments, the Cathepsin K polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 10. In some embodiments, the Cathepsin K polypeptide comprises the amino acid sequence of SEQ ID NO: 10.

[0082] In some embodiments, the at least one catabolic enzyme is Cathepsin L (a.k.a. MEP, CTSL, EC 3.4.22.15, CATL, Major Excreted Protein). Cathepsin L is a lysosomal endopeptidase enzyme which is involved in the initiation of protein degradation. Its substrates include collagen and elastin, as well as alpha-1 protease inhibitor, a major controlling element of neutrophil elastase activity. In some embodiments, the Cathepsin L can be a natural, synthetic, or recombinant protein. In some embodiments, the Cathepsin L polypeptide comprises an amino acid sequence with at least about 70%, 71%, 72%, 73%, 74%, 75%, 76%, 77%, 78%, 79%, 80%, 81%, 82%, 83%, 84%, 85%, 86%, 87%, 88%, 89%, 90%, 91%, 92%, 93%, 94%, 95%, 96%, 97%, 98%, 99%, or more sequence identity to SEQ ID NO: 12, 59, 61, 63, 65, or 67. In some embodiments, the Cathepsin L polypeptide comprises the amino acid sequence of SEQ ID NO: 12, 59, 61, 63, 65, or 67.

[0083] In some embodiments, the administration comprises the administration of a nucleotide sequence encoding at least one catabolic enzyme of the present invention.

[0084] As used herein, the terms "polynucleotide", "polynucleotide sequence", "nucleic acid sequence", "nucleic acid fragment", "nucleotide sequence," and "isolated nucleic acid fragment" are used interchangeably herein. These terms encompass nucleotide sequences and the like. A polynucleotide may be a polymer of RNA or DNA that is single- or double-stranded, that optionally contains synthetic, non-natural or altered nucleotide bases. A polynucleotide in the form of a polymer of DNA may be comprised of one or more segments of cDNA, genomic DNA, synthetic DNA, or mixtures thereof. Nucleotides (usually found in their 5'-monophosphate form) are referred to by a single letter designation as follows: "A" for adenylate or deoxyadenylate (for RNA or DNA, respectively), "C" for cytidylate or deoxycytidylate, "G" for guanylate or deoxyguanylate, "U" for uridylate, "T" for deoxythymidylate, "R" for purines (A or G), "Y" for pyrimidines (C or T), "K" for G or T, "H" for A or C or T, "I" for inosine, and "N" for any nucleotide.

[0085] As used herein, the term "chimeric" or "recombinant" when describing a nucleic acid sequence or a protein sequence refers to a nucleic acid or a protein sequence that links at least two heterologous polynucleotides or two heterologous polypeptides into a single macromolecule, or that re-arranges one or more elements of at least one natural nucleic acid or protein sequence. For example, the term "recombinant" can refer to an artificial combination of two otherwise separated segments of sequence, e.g., by chemical synthesis or by the manipulation of isolated segments of nucleic acids by genetic engineering techniques.

[0086] As used herein, a "synthetic nucleotide sequence" or "synthetic polynucleotide sequence" is a nucleotide sequence that is not known to occur in nature or that is not naturally occurring. Generally, such a synthetic nucleotide sequence will comprise at least one nucleotide difference when compared to any other naturally occurring nucleotide sequence. It is recognized that a genetic regulatory element of the present invention comprises a synthetic nucleotide sequence. In some embodiments, the synthetic nucleotide sequence shares little or no extended homology to natural sequences. Extended homology in this context generally refers to 100% sequence identity extending beyond about 25 nucleotides of contiguous sequence. A synthetic genetic regulatory element of the present invention comprises a synthetic nucleotide sequence.

[0087] As used herein, an "isolated" or "purified" nucleic acid molecule or polynucleotide, or biologically active portion thereof, is substantially or essentially free from components that normally accompany or interact with the

nucleic acid molecule or polynucleotide as found in its naturally occurring environment. Thus, an isolated or purified nucleic acid molecule or polynucleotide is substantially free of other cellular material or culture medium when produced by recombinant techniques, or substantially free of chemical precursors or other chemicals when chemically synthesized.

[0088] In some embodiments, the methods comprise administering to the subject a composition comprising an expression vector (interchangeably referred to herein as a vector), wherein the vector comprises a polynucleotide sequence encoding at least one catabolic enzyme. In some embodiments, the methods comprise administering to the subject a composition comprising at least one expression vector comprising an expression cassette of coding genes.

[0089] In some embodiments, the expression vector is a viral vector. Accordingly, in the some embodiments, the methods of the present invention comprise administering to the subject a composition comprising at least one viral vector comprising a polynucleotide sequence encoding at least one catabolic enzyme. In some embodiments, the expression cassette, the expression vector, or the viral vector further comprises one or more nucleotide sequences encoding a signal peptide. In some embodiments, the signal peptide is an intralysosomal localization peptide.

[0090] A nucleotide sequence encoding at least one catabolic enzyme can be delivered to a subject through any suitable delivery system, such as those described by Rolland (Pharmaceutical Gene Delivery Systems, ISBN: 978-0-8247-4235-5, 2003), which is incorporated by reference in its entirety. In some embodiments, the delivery system is a viral system, a physical system, and/or a chemical system.

[0091] In some embodiments, the delivery system to deliver a nucleotide sequence encoding at least one catabolic enzyme is a viral system. In some embodiments, an adenovirus vector is used (see, Thrasher et al., Gene therapy: X-SCID transgene leukaemologenicity. Nature. 2006; 443 (7109): E5-E6; Zhang et al., Adenoviral and adeno-associated viral vectors-mediated neuronal gene transfer to cardiovascular control regions of the rat brain. Int J Med Sci. 2013; 10(5): 607-616). In some embodiments, an adeno-associated vector is used (see, Teramato et al., Crisis of adenoviruses in human gene therapy. Lancet. 2000; 355 (9218): 1911-1912, Okada et al., Gene transfer targeting mouse vestibule using adenovirus and adeno-associated virus vectors. Otol Neurotol. 2012; 33(4): 655-659). In some embodiments, a retroviral vector is used (see, Anson et al., The use of retroviral vectors for gene therapy-what are the risks? A review of retroviral pathogenesis and its relevance to retroviral vector-mediated gene delivery. Genet Vaccines Ther. 2004; 2(1): 9; Frederic D. Retroviral integration and human gene therapy. J Clin Invest. 2007; 117(8): 2083-2086). In some embodiments, a lentivirus vector is used (see, Goss et al., Antinociceptive effect of a genomic herpes simplex virus-based vector expressing human proenkephalin in rat dorsal root ganglion. Gene Ther. 2001; 8(7): 551-556; Real et al., Improvement of lentiviral transfer vectors using cis-acting regulatory elements for increased gene expression. Appl Microbiol Biotechnol. 2011; 91(6): 1581-91.). In some embodiments, a herpes simplex virus vector is used (see, Lachmann R H, Efstathiou S. The use of herpes simplex virus-based vectors for gene delivery to the nervous system. Mol Med Today. 1997; 3(9): 404-411; Liu S, Dai M, You L, Zhao Y. Advance in herpes simplex viruses for cancer therapy. Sci China Life Sci. 2013; 56(4): 298-305). In some embodiments, a poxvirus vector is used (see, Moss B. Reflections on the early development of poxvirus vectors. Vaccine. 2013; 31(39): 4220-4222). Each of the references is incorporated herein by reference in its entirety.

[0092] In some embodiments, the delivery system to deliver a nucleotide sequence encoding at least one catabolic enzyme of the invention is a physical system. In some embodiments, the physical systems include, but are not limited to jet injection, biolistics, electroporation, hydrodynamic injection, and ultrasound (see, Sirsi et al. Advances in ultrasound mediated gene therapy using microbubble contrast agents. Theranostics. 2012; 2(12): 1208-1222.; Naldini et al., In vivo gene delivery and stable transduction of nondividing cells by a lentiviral vector. Science. 1996; 272(5259): 263-267; Panje et al., Ultrasound-mediated gene delivery with cationic versus neutral microbubbles: Effect of DNA and microbubble dose on in vivo transfection efficiency. Theranostics. 2012; 2(11): 1078-1091; Gao et al., Cationic liposome-mediated gene transfer. Gene Ther. 1995; 2(10): 710-722; Orio et al., Electric field orientation for gene delivery using high-voltage and low-voltage pulses. J Membr Biol. 2012; 245(10): 661-666.) Each of the references is incorporated herein by reference in its entirety.

[0093] In some embodiments, the delivery system to deliver a nucleotide sequence encoding at least one catabolic enzyme of the invention is a chemical system. The chemical systems include, but are not limited to calcium phosphate precipitation, liposomes and polymeric carriers. In some embodiments, the chemical system is based on calcium phosphate precipitation, such as calcium phosphate nanocomposite particles encapsulating DNA (see, Nouri et al. Calcium phosphate-mediated gene delivery using simulated body fluid (SBF). Int J Pharm. 2012; 434(1-2): 199-208; Bhakta et al. Magnesium phosphate nanoparticles can be efficiently used in vitro and in vivo as non-viral vectors for targeted gene delivery. J Biomed Nanotechnol. 2009; 5(1): 106-114).

[0094] In some embodiments, the chemical system to deliver a nucleotide sequence encoding at least one catabolic enzyme of the invention is based on liposomes. In some embodiments, the liposomes are nano-sized. In some embodiments, liposomes conjugated with polyethylene glycol (PEG) and/or other molecules such as ligands and peptides can be used (see, Yang et al. Cationic nucleolipids as efficient siRNA carriers. Org Biomol Chem. 2011; 1(9): 291-296).

[0095] In some embodiments, the chemical system to deliver a nucleotide sequence encoding at least one catabolic enzyme of the invention is based on polymeric carriers. In some embodiments, the polymeric carriers are conjugated to the gene to be delivered. In some embodiments, the polymeric carriers include, but are not limited to chitosan, polyethylenimine (PEI), polylysine, polyarginine, polyamino ester, Polyamidoamine Dendrimers (PAMAM), Poly (lactide-co-glycolide), and PLL, such as those described in Choi et al., Enhanced transfection efficiency of PAMAM dendrimer by surface modification with 1-arginine. J Control Release. 2004; 3(99): 445-456; Pfeifer et al., Poly(ester-anhydride):poly(beta-amino ester) micro- and nanospheres: DNA encapsulation and cellular transfection. Int J Pharm. 2005; 304(1-2): 210-219; Anderson et al., Structure/property studies of polymeric gene delivery using a library of poly(beta-amino esters). Mol Ther. 2005; 3(11):

426-434; Hwang et al., Effects of structure of beta-cyclodextrin-containing polymers on gene delivery. Bioconjugate Chem. 2001; 2(12): 280-290; Kean et al., Trimethylated chitosans as non-viral gene delivery vectors: cytotoxicity and transfection efficiency. J Control Release. 2005; 3(103): 643-653.

[0096] In some embodiments, administration of a catabolic enzyme comprises the administration of at least one catabolic enzyme polypeptide or fragment thereof of the present invention. As used herein, the terms "polypeptide" and "protein" are used interchangeably herein.

[0097] The invention also envisions and encompasses the use of functional variants or fragments of the intralysosomal catabolic enzyme described herein. As used herein, the phrase "a biologically active variant" or "functional variant" with respect to a protein refers to an amino acid sequence that is altered by one or more amino acids with respect to a reference sequence, while still maintains substantial biological activity of the reference sequence. The variant can have "conservative" changes, wherein a substituted amino acid has similar structural or chemical properties, e.g., replacement of leucine with isoleucine. The following table shows exemplary conservative amino acid substitutions.

Original Residue	Very Highly - Conserved Substitutions	Highly Conserved Substitutions (from the Blosum90 Matrix)	Conserved Substitutions (from the Blosum65 Matrix)
Ala	Ser	Gly, Ser, Thr	Cys, Gly, Ser, Thr, Val
Arg	Lys	Gln, His, Lys	Asn, Gln, Glu, His, Lys
Asn	Gln; His	Asp, Gln, His, Lys, Ser, Thr	Arg, Asp, Gln, Glu, His, Lys, Ser, Thr
Asp	Glu	Asn, Glu	Asn, Gln, Glu, Ser
Cys	Ser	None	Ala
Gln	Asn	Arg, Asn, Glu, His, Lys, Met	Arg, Asn, Asp, Glu, His, Lys, Met, Ser
Glu	Asp	Asp, Gln, Lys	Arg, Asn, Asp, Gln, His, Lys, Ser
Gly	Pro	Ala	Ala, Ser
His	Asn; Gln	Arg, Asn, Gln, Tyr	Arg, Asn, Gln, Glu, Tyr
Ile	Leu; Val	Leu, Met, Val	Leu, Met, Phe, Val
Leu	Ile; Val	Ile, Met, Phe, Val	Ile, Met, Phe, Val
Lys	Arg; Gln; Glu	Arg, Asn, Gln, Glu	Arg, Asn, Gln, Glu, Ser,
Met	Leu; Ile	Gln, Ile, Leu, Val	Gln, Ile, Leu, Phe, Val
Phe	Met; Leu; Tyr	Leu, Trp, Tyr	Ile, Leu, Met, Trp, Tyr
Ser	Thr	Ala, Asn, Thr	Ala, Asn, Asp, Gln, Glu, Gly, Lys, Thr
Thr	Ser	Ala, Asn, Ser	Ala, Asn, Ser, Val
Trp	Tyr	Phe, Tyr	Phe, Tyr
Tyr	Trp; Phe	His, Phe, Trp	His, Phe, Trp
Val	Ile; Leu	Ile, Leu, Met	Ala, Ile, Leu, Met, Thr

[0098] Alternatively, a variant can have "nonconservative" changes, e.g., replacement of a glycine with a tryptophan. Analogous minor variations can also include amino acid deletion or insertion, or both. Guidance in determining which amino acid residues can be substituted, inserted, or deleted without eliminating biological or immunological activity can be found using computer programs well known in the art, for example, DNASTAR software. For polynucleotides, a variant comprises a polynucleotide having deletions (i.e., truncations) at the 5' and/or 3' end; deletion and/or addition of one or more nucleotides at one or more internal sites in the reference polynucleotide; and/or substitution of one or more nucleotides at one or more sites in the reference polynucleotide. As used herein, a "reference" polynucleotide comprises a nucleotide sequence produced by the methods disclosed herein. Variant polynucleotides also include synthetically derived polynucleotides, such as those generated, for example, by using site directed mutagenesis but which still comprise genetic regulatory element activity. Generally, variants of a particular polynucleotide or nucleic acid molecule, or polypeptide of the invention will have at least about 60%, 65%, 70%, 75%, 80%, 85%, 90%, 91%, 91.5%, 92%, 92.5%, 93%, 93.5%, 94%, 94.5%, 95%, 95.5%, 96%, 96.5%, 97%, 97.5%, 98%, 98.5%, 99%, 99.1%, 99.2%, 99.3%, 99.4%, 99.5%, 99.6%, 99.7%, 99.8%, 99.9% or more sequence identity to that particular polynucleotide/polypeptides as determined by sequence alignment programs and parameters as described elsewhere herein.

[0099] In some embodiments, a gene that can hybridize with the nucleic acid sequences encoding the catabolic enzymes of the present invention under stringent hybridization conditions can be used. The terms "stringency" or "stringent hybridization conditions" refer to hybridization conditions that affect the stability of hybrids, e.g., temperature, salt concentration, pH, formamide concentration and the like. These conditions are empirically optimized to maximize specific binding and minimize non-specific binding of primer or probe to its target nucleic acid sequence. The terms as used include reference to conditions under which a probe or primer will hybridize to its target sequence, to a detectably greater degree than other sequences (e.g. at least 2-fold over background). Stringent conditions are sequence dependent and will be different in different circumstances. Longer sequences hybridize specifically at higher temperatures. Generally, stringent conditions are selected to be about 5° C. lower than the thermal melting point (Tm) for the specific sequence at a defined ionic strength and pH. The Tm is the temperature (under defined ionic strength and pH) at which 50% of a complementary target sequence hybridizes to a perfectly matched probe or primer. Typically, stringent conditions will be those in which the salt concentration is less than about 1.0 M Na^+ ion, typically about 0.01 to 1.0 M Na+ion concentration (or other salts) at pH 7.0 to 8.3 and the temperature is at least about 30° C. for short probes or primers (e.g. 10 to 50 nucleotides) and at least about 60° C. for long probes or primers (e.g. greater than 50 nucleotides). Stringent conditions may also be achieved with the addition of destabilizing agents such as formamide. Exemplary low stringent conditions or "conditions of reduced stringency" include hybridization with a

buffer solution of 30% formamide, 1 M NaCl, 1% SDS at 37° C. and a wash in 2×SSC at 40° C. Exemplary high stringency conditions include hybridization in 50% formamide, 1M NaCl, 1% SDS at 37° C., and a wash in 0.1×SSC at 60° C. Hybridization procedures are well known in the art and are described by e.g. Ausubel et al., 1998 and Sambrook et al., 2001. In some embodiments, stringent conditions are hybridization in 0.25 M Na_2HPO_4 buffer (pH 7.2) containing 1 mM Na_2EDTA, 0.5-20% sodium dodecyl sulfate at 45° C., such as 0.5%, 1%, 2%, 3%, 4%, 5%, 6%, 7%, 8%, 9%, 10%, 11%, 12%, 13%, 14%, 15%, 16%, 17%, 18%, 19% or 20%, followed by a wash in 5×SSC, containing 0.1% (w/v) sodium dodecyl sulfate, at 55° C. to 65° C.

[0100] The definition of each catabolic enzyme includes sequences having high similarity or identity to the nucleic acid sequences and/or polypeptide sequences of the specific catabolic enzymes mentioned herein. As used herein, "sequence identity" or "identity" in the context of two nucleic acid or polypeptide sequences includes reference to the residues in the two sequences which are the same when aligned for maximum correspondence over a specified comparison window. When percentage of sequence identity is used in reference to proteins it is recognized that residue positions which are not identical often differ by conservative amino acid substitutions, where amino acid residues are substituted for other amino acid residues with similar chemical properties (e.g., charge or hydrophobicity) and therefore do not change the functional properties of the molecule. Where sequences differ in conservative substitutions, the percent sequence identity may be adjusted upwards to correct for the conservative nature of the substitution. Sequences which differ by such conservative substitutions are said to have "sequence similarity" or "similarity." Means for making this adjustment are well-known to those of skill in the art. Typically this involves scoring a conservative substitution as a partial rather than a full mismatch, thereby increasing the percentage sequence identity. Thus, for example, where an identical amino acid is given a score of 1 and a non-conservative substitution is given a score of zero, a conservative substitution is given a score between zero and 1. The scoring of conservative substitutions is calculated, e.g., according to the algorithm of Meyers and Miller, Computer Applic. Biol. Sci., 4:11-17 (1988).

[0101] The invention also includes biologically active fragments of the catabolic enzymes described herein. These biologically active fragments may comprise at least 10, 20, 50, 100, 150, 200, 250, 300, 350, 400, 450, or more amino acid residues and retain one or more activities associated with the catabolic enzymes described herein. Such fragments may be obtained by deletion mutation, by recombinant techniques that are routine and well-known in the art, or by enzymatic digestion of the catabolic enzyme(s) of interest using any of a number of well-known proteolytic enzymes. The invention further includes nucleic acid molecules which encode the above described variant enzymes and enzyme fragments.

[0102] In some embodiments, the methods comprise administering to the subject a composition comprising a therapeutically effective amount or prophylactically effective amount of at least one catabolic enzyme. The term "therapeutically effective amount" as used herein, refers to the level or amount of one or more catabolic enzymes needed to treat amyloidosis, or reduce or prevent injury or damage, optionally without causing significant negative or adverse side effects. A "prophylactically effective amount" refers to an amount of a catabolic enzyme sufficient to prevent or reduce severity of a future disease or condition associated with amyloidosis when administered to a subject who is susceptible and/or who may develop amyloidosis or a condition associated with amyloidosis.

[0103] In some embodiments, instead of or in addition to administering a polynucleotide sequence encoding a catabolic enzyme of the present invention, the methods comprise administering a composition comprising a polypeptide comprising a catabolic enzyme of the present invention or a biologically active fragment thereof directly to the subject in need.

[0104] In some embodiments, the catabolic enzyme is targeted to the intralysosomal space. In some embodiments, the catabolic enzyme to be administered comprises one or more signals which help with sorting the polypeptide to lysosome. In some embodiments, the signal can be a lysosomal localization signal polypeptide, a monosaccharide (including derivatives), a polysaccharide, or combinations thereof.

[0105] In some embodiments, the signal is mannose-6 phosphate. A catabolic enzyme comprising a mannose-6 phosphate can be targeted to lysosomes with the help of a mannose-6 phosphate receptor.

[0106] In some embodiments, the signal is not dependent on mannose-6 phosphate. In some embodiments, the signal is a signal peptide. In some embodiments, the signal peptide is located at the N-terminal, the C-terminal, or elsewhere in the intralysosomal catabolic enzyme to be administered. In some embodiments, the signal peptides include, but are not limited to the DXXLL type (SEQ ID NO: 13), [DE]XXXL[LI] type (SEQ ID NO: 14), and YXXO type (SEQ ID NO: 15). See Bonifacino et al., Signals for sorting of transmembrane proteins to endosomes and lysosomes, Annu. Rev. Biochem. 72 (2003) 395-447; and Brualke et al. (Sorting of lysosomal proteins, Biochimica et Biophysica Acta 1793 (2009) 605-614), each of which is incorporated by reference in its entirety.

[0107] In some embodiments, the signal peptides belong to the DXXLL type, such as those identified in MPR300/CI-MPR (SFHDDS**DEDLL**, SEQ ID NO: 16), MPR46/CD-MPR (EESEER**DDHLL**, SEQ ID NO: 17), Sortilin (GYHDDS**DEDLL**, SEQ ID NO: 18), SorLA/SORL1 (ITGFSD**DVPMV**, SEQ ID NO: 19), GGA1 (1) (ASVSLL**DDELM**, SEQ ID NO: 20), GGA1 (2) (ASSGLD**DLDLL**, SEQ ID NO: 21), GGA2 (VQNPSA**DRNLL**, SEQ ID NO: 22), and GGA3 (NALSWL**DEELL**, SEQ ID NO: 23).

[0108] In some embodiments, the signal peptides belong to the [DE]XXXL[LI] type, such as those identified in LIMP-II (**DERAPLI**, SEQ ID NO: 24), NPC1 (**TERERLL**, SEQ ID NO: 25), Mucolipin-1 (**SETERLL**, SEQ ID NO: 26), Sialin (**TDRTPLL**, SEQ ID NO: 27), GLUT8 (**EETQPLL**, SEQ ID NO: 28), Invariant chain (Ii) (1) (**DDQRDLI**, SEQ ID NO: 29), and Invariant chain (Ii) (2) (**NEQLPML**, SEQ ID NO: 30).

[0109] In some embodiments, the signal peptides belong to the YXXO type, such as those identified in LAMP-1 (**GYQTI**, SEQ ID NO: 31), LAMP-2A (**GYEQF**, SEQ ID NO: 32), LAMP-2B (**GYQTL**, SEQ ID NO: 33), LAMP-2C (**GYQSV**, SEQ ID NO: 34), CD63 (**GYEVM**, SEQ ID NO: 35), CD68 (**AYQAL**, SEQ ID NO: 36), Endolyn (**NYHTL**, SEQ ID NO: 37), DC-LAMP (**GYQRI**, SEQ ID

NO: 38), Cystinosin (GYDQL, SEQ ID NO: 39), Sugar phosphate exchanger 2 (GYKEI, SEQ ID NO: 40), and acid phosphatase (GYRHV, SEQ ID NO: 41).

[0110] In some embodiments, the catabolic enzyme is targeted to remain outside the cell, i.e., the enzyme is modified to act extracellularly. In some embodiments, the catabolic enzyme to be administered lacks one or more signals that would otherwise target the polypeptide to the lysosome. In some embodiments, the catabolic enzyme lacks one or more mannose-6 phosphate (i.e., M6P) signals, thereby precluding entry of the catabolic enzyme into the cell. In some embodiments, the catabolic enzyme is recombinantly engineered to lack one or more mannose-6 phosphate signal. Not bound by any theory, it is generally understood in the art that reduced M6P content lowers the binding affinity of a recombinant enzyme for M6P receptors and decreases its cellular uptake and thereby allows the enzyme to remain outside the cell.

[0111] Methods for reducing the M6P content of a recombinant protein, e.g., a catabolic enzyme, are known in the art. See, e.g., U.S. Pat. No. 8,354,105, which is herein incorporated by reference in its entirety. In some embodiments, the mannose content of a recombinant catabolic enzyme may be reduced by manipulating the cell culture conditions such that the glycoprotein produced by the cell has low-mannose content. As used herein, the term "low-mannose content" refers to catabolic enzyme composition wherein less than about 20%, less than about 15%, less than about 10%, less than about 8%, less than about 5%, less than about 4%, less than about 3%, less than about 2%, less than about 1%, or any values between any of these preceding ranges, or even at 0% of the enzymes in the composition have more than 4 mannose residues (i.e.. are species of M5 or greater).

[0112] In some embodiments, the present invention provides a composition comprising at least two catabolic enzymes, wherein the composition comprises at least one catabolic enzyme that is targeted to the cell lysosome and at least one catabolic enzyme that remains outside the cell. In some embodiments, the catabolic enzymes are selected from protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L. In an exemplary embodiment, the present invention provides a composition comprising at least two catabolic enzymes, wherein the composition comprises a PPCA catabolic enzyme that is targeted to the cell lysosome and a PPCA catabolic enzyme that remains outside the cell. In some embodiments, the ratio of the intralysosomal catabolic enzyme to the extracellular catabolic enzyme on a percentage basis within the composition is at least 5%:95%. In further embodiments, the ratio of the intralysosomal catabolic enzyme to the extracellular catabolic enzyme on a percentage basis within the composition is at least 10%:90%, at least 15%:85%, at least 20%:80%, at least 25%:75%, at least 30%:70%, at least 35%:65%, at least 40%:60%, at least 45%:55%, at least 50%:50%, at least 55%:45%, at least 60%:40%, at least 65%:35%, at least 70%:30%, at least 75%:25%, at least 80%:20%, at least 85%:15%, at least 90%:10%, or at least 95%:5%.

[0113] In some embodiments, the methods of the present invention comprise administering to the subject a composition comprising a therapeutically effective amount of at least two, three, or more catabolic enzymes. In some embodiments, the methods comprise increasing the expression, activity, and/or concentration of at least two, three, or more catabolic enzymes in the subject. In some embodiments, the methods comprise administering to the subject a composition comprising an expression cassette comprising one or more polynucleotide sequences encoding at least two, three, or more catabolic enzymes. In some embodiments, the methods comprise administering to the subject one or more expression cassettes comprising at least two, three or more polynucleotide sequences encoding at least two, three or more catabolic enzymes. In some embodiments, the methods comprise administering to the subject a therapeutically effective amount of a first catabolic enzyme, and an expression cassette comprising a polynucleotide sequence encoding a second catabolic enzyme. In some embodiments, two or more catabolic enzymes are selected from the group consisting of protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L. In some embodiments, at least two catabolic enzymes are PPCA and NEU1.

[0114] In some embodiments, administration of the at least one catabolic enzyme is employed to prevent the formation of amyloid. In other embodiments, administration of the at least one catabolic enzyme is employed to degrade amyloid that has already formed. In some embodiments, administration of the at least one catabolic enzyme is employed to prevent the formation of one or more amyloid oligomers. In some embodiments, administration of the at least one catabolic enzyme is employed to prevent the formation of one or more amyloid fibrils. In some embodiments, administration of the at least one catabolic enzyme is employed to degrade one or more amyloid oligomers after it has already formed. In some embodiments, administration of the at least one catabolic enzyme is employed to degrade one or more amyloid fibrils after it has already formed.

[0115] In some embodiments, the methods of the present invention provided herein further comprise administering a composition (e.g. a pharmaceutical composition) comprising at least one catabolic enzyme or fragment thereof with at least one additional drug for treating or preventing amyloidosis.

[0116] In some embodiments, the at least one additional drug is a steroid. In some embodiments, the steroid is dexamethasone, cortisone, hydrocortisone, methylprednisolone, prednisolone, prednisone, triamcinolone or any combination thereof.

[0117] In some embodiments, the at least one additional drug is a non-steroid agent. In some embodiments, such non-steroid agent is diclofenac, flufenamic acid, flurbiprofen, diflunisal, detoprofen, diclofenac, etodolac, fenoprofen, ibuprofen, indomethacin, ketoprofen, meclofenameate, mefenamic acid, meloxicam, nabumeone, naproxen sodium, oxaprozin, piroxicam, sulindac, tolmetin, celecoxib, rofecoxib, aspirin, choline salicylate, salsalte, and sodium and magnesium salicylate or any combination thereof.

[0118] In some embodiments, the at least one additional drug is a chemotherapy agent. In some embodiments, the chemotherapy agent is selected from the group consisting of cyclophosphamide (e.g., Cytoxan, Neosar) and melphalan (e.g., Alkeran).

[0119] In some embodiments, at least one additional drug is an anti-inflammatory medication, when the subject has inflammatory symptoms.

[0120] In some embodiments, the at least one additional drug is an antibiotic, when the subject has infection symptoms. In some embodiments, the infection is a chromic infection. In some embodiments, the infection is a microbial infection.

[0121] In some embodiments, the at least one additional drug is a Carbonic Anhydrase (CA) enzyme (e.g., CA-I, CA-II, CA-III, CA-IV, CA-V, CA-VI, and CA-VII) and/or agents that can increase the activity of a Carbonic Anhydrase enzyme in the subject.

[0122] In some embodiments, at least one additional drug is a disease modifying antirheumatic drug (DMARD). In some embodiments, the DMARD is cyclosporine, azathioprine, methotrexate, leflunomide, cyclophosphamide, hydroxychloroquine, sulfasalazine, D-penicillamine, minocycline, gold, or any combination thereof.

[0123] In some embodiments, the at least one additional drug is a recombinant protein. In some embodiments, the recombinant protein is ENBREL® (etanercept, a soluble TNF receptor) or REMICADE® (infliximab, a chimeric monoclonal anti-TNF antibody).

[0124] In some embodiments, the one or more additional drugs is/are selected from melphalan, dexamethasone, bortezomib, lenalidomide, vincristine, doxorubicin, cyclophosphamide and pomalidomide.

[0125] In some embodiments, the methods of the present invention further comprise the administration of one or more drugs that acidifies the lysosome. As used herein, drugs that acidify the lysosome are drugs capable of lowering the lysosomal pH of a target cell. Accordingly, in some embodiments, the present invention provides a method of treating or preventing amyloidosis in a subject comprising administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof, wherein the subject is also administered one or more drugs that acidifies the lysosome. As described herein, when performing a combination therapy, the two or more drugs (e.g., a catabolic enzyme or a biologically active fragment thereof and a drug that acidifies the lysosome) can be administered simultaneously or sequentially in any order.

[0126] In some embodiments, the drug that acidifies the lysosome is selected from an acidic nanoparticle, a catecholamine, a β-adrenergic receptor agonist, an adenosine receptor agonist, a dopamine receptor agonist, an activator of the cystic fibrosis transmembrane conductance regulator (CFTR), cyclic adenosine monophosphate (cAMP), a cAMP analog, and an inhibitor of glycogen synthase kinase-3 (GSK-3).

[0127] In some embodiments, the drug that acidifies the lysosome is an acidic nanoparticle. Acidic nanoparticles have been shown to localize to lysosomes and reduce lysosomal pH. See Baltazar et al., 2012, PloS ONE 7(12): e49635 and Lee et al., 2015, *Cell Rep.* 12(9): 1430-44, both of which are herein incorporated by reference in their entireties. In some embodiments, the acidic nanoparticle is a polymeric acidic nanoparticle. In some embodiments, the polymeric acidic nanoparticle is a poly (DL-lactide-co-glycolide) (PLGA) acidic nanoparticle. In a specific embodiment, the PLGA acidic nanoparticle comprises PLGA Resomer RG 503 H. In some embodiments, the PLGA acidic nanoparticle comprises PLGA Resomer RG 502 H. In other embodiments, the polymeric acidic nanoparticle is a poly (DL-lactide) (PLA) acidic nanoparticle. In a specific embodiment, the PLA acidic nanoparticle comprises PLA Resomer R 203 S. In some embodiments, the acid number of the acidic nanoparticle is between about 0.5 mg KOH/g to about 8 mg KOH/g. In some embodiments, the acid number of the acidic nanoparticle is between about 1 mg KOH/g to about 6 mg KOH/g. In some embodiments, the acid number of the acidic nanoparticle is selected from about 1 mg KOH/g, about 2 mg KOH/g, about 3 mg KOH/g, about 4 mg KOH/g, about 5 mg KOH/g, or about 6 mg KOH/g. In a specific embodiment, the acid number of the acidic nanoparticle is about 3 mg KOH/g. In some embodiments, the nanoparticle size is about 50 nm to about 800 nm. In some embodiments, the nanoparticle size is about 100 nm to about 600 nm. In a specific embodiment, the nanoparticle size is about 350 nm to about 550 nm. In a further specific embodiment, the nanoparticle size is about 375 nm to about 400 nm. In an exemplary embodiment, the acidic nanoparticle is spherical. In some embodiments, the nanoparticles are targeting a specific transport process in the brain, which enhance drug transport through the blood-brain barrier (BBB). In some embodiments, such transport processes include, but are not limited to: (1) nanoparticles open TJs between endothelial cells or induce local toxic effect which leads to a localized permeabilization of the BBB allowing the penetration of the drug in a free form or conjugated with the nanoparticles; (2) nanoparticles pass through endothelial cell by transcytosis; (3) nanoparticles are transported through endothelial cells by endocytosis, where the content is released into the cell cytoplasm and then exocytosed in the endothelium abluminal side; and (4) a combination of several of the mechanisms. In some embodiments, the receptors targeted by nanoparticles are transferrin and low-density lipo-protein receptors. In some embodiments, the targeting can be achieved by peptides, proteins, or antibodies, which can be physically and/or chemically immobilized on the nanoparticles. In some embodiments, the nanoparticles are coated with one or more apolipoproteins, such as apolipoprotein AII, B, CII, E, and/or J (see, Kreuter et al., (2002, DOI: 10.1080/10611860290031877). For more nanoparticle-mediated brain drug delivery compositions and methods, see Saraiva et al. (Journal of Controlled Release, 2016, 235:34-37). Each of the references mentioned herein is incorporated by reference in its entirety.

[0128] In some embodiments, the drug that acidifies the lysosome is a catecholamine. Catecholamines have been shown to reduce lysosomal pH. See Liu et al., 2008, *Invest Ophthalmol Vis Sci.* 49(2): 772-780, which is herein incorporated by reference in its entirety. In some embodiments, the catecholamine is selected from epinephrine, metanephrine, synephrine, norepinephrine, normetanephrine, octopamine or norphenephrine, dopamine, and dopa. In exemplary embodiment, the catecholamine is selected from epinephrine, norepinephrine, and dopamine.

[0129] In some embodiments, the drug that acidifies the lysosome is a β-adrenergic receptor agonist. β-adrenergic receptor agonists have been shown to reduce lysosomal pH. See Liu et al., 2008, Invest Ophthalmol Vis Sci. 49(2): 772-780. Examples of β-adrenergic receptor agonists may be found in US Patent Publication No. 2012/0329879, which is herein incorporated by reference in its entirety. In some embodiments, the β-adrenergic receptor agonist is selected from isoproterenol, metaproterenol, formoterol, salmeterol,

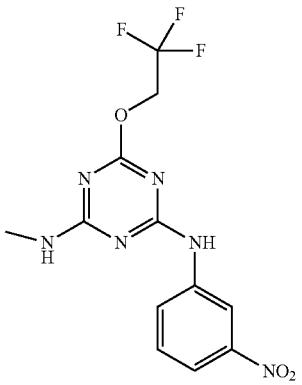
salbutamol, albuterol, terbutaline, fenoterol, and vilanterol. In an exemplary embodiment, the β-adrenergic receptor agonist is isoproterenol.

[0130] In some embodiments, the drug that acidifies the lysosome is an adenosine receptor agonist. Adenosine receptor agonists have been shown to reduce lysosomal pH. See Liu et al., 2008, Invest Ophthalmol Vis Sci. 49(2): 772-780. In an exemplary embodiment, the adenosine receptor agonist is a non-specific adenosine receptor agonist or an A_{2A} adenosine receptor agonist. Examples of A_{2A} adenosine receptor agonists may be found in US Patent Publication No. 2012/0130481, which is herein incorporated by reference in its entirety. In some embodiments, the adenosine receptor agonist is selected from 5'-N-ethylcarboxamidoadenosine (NECA), CGS21680, 2-phenylaminoadenosine, 2-[para-(2carboxyethyl)phenyl]amino-5'N-ethylcarboxamidoadenosine, SRA-082, 5'-N-cyclopropylcarboxamidoadenosine, 5'N-methylcarboxamidoadenosine and PD-125944.

[0131] In some embodiments, the drug that acidifies the lysosome is a dopamine receptor agonist. Dopamine receptor agonists have been shown to reduce lysosomal pH. See Guha et al., 2014, *Adv Exp Med Biol.* 801: 105-111, which is herein incorporated by reference in its entirety. In some embodiments, the dopamine receptor agonist is selected from A68930, A77636, A86929, SKF81297, SKF82958, SKF38393, SKF89145, SKF89626, dihydrexidine, dinapsoline, dinoxyline, doxanthrine, fenoldopam, 6-Br-APB, stepholidine, CY-208243, 7,8-Dihydroxy-5-phenyl-octahydrobenzo[h]isoquinoline, cabergoline, and pergolide. In an exemplary embodiment, the dopamine receptor agonist is selected from A68930, A77636, and SKF81297. In a further exemplary embodiment, the dopamine receptor agonist is SKF81297, also known as 6-chloro-1-phenyl-2,3,4,5-tetrahydro-1H-3-benzazepine-7,8-diol.

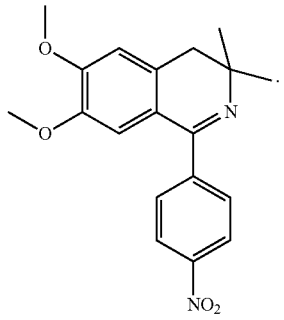
[0132] In some embodiments, the drug that acidifies the lysosome is an activator of the cystic fibrosis transmembrane conductance regulator (CFTR). Activators of CFTR have been shown to reduce lysosomal pH. See Liu et al., 2012, *Am J Physiol Cell Physiol* 303: C160-9, which is herein incorporated by reference in its entirety. In some embodiments, the CFTR activator is selected from $CFTR_{Act}01$ to $CFTR_{Act}17$. See Ma et al., *J Biol Chem* 277: 37235-37241. In an exemplary embodiment, the CFTR activator is selected from $CFTR_{Act}11$ and $CFTR_{Act}16$, having the following structures:

$CFTR_{Act}11$



-continued

$CFTR_{Act}16$



In some embodiments, the CFTR activator is co-administered with forskolin.

[0133] In some embodiments, the drug that acidifies the lysosome is cAMP or a cAMP analog. cAMP and/or cAMP analogs have been shown to reduce lysosomal pH. See Liu et al., 2008, Invest Ophthalmol Vis Sci. 49(2): 772-780. For instance, the cell-permeable analogs chlorophenylthio-cAMP (cpt-cAMP) and 8-bromo-cAMP have the ability to lower lysosomal pH in cells. In some embodiments, cAMP and/or a cAMP analog may be administered in a cocktail comprising 3-isobutyl-1-methylxanthine (IBMX) and forskolin. For example, in one embodiment, a cocktail comprising IBMX, forskolin, and cpt-cAMP may be administered to acidify the lysosome. In some embodiments, the cAMP analog is selected from 9-pCPT-2-O-Me-cAMP, Rp-cAMPS, 8-Cl-cAMP, Dibutyryl cAMP, pCPT-cAMP, N6-monobutyryladenosine 3',5'-cyclic monophosphate, and PDE inhibitors.

[0134] In some embodiments, the drug that acidifies the lysosome is an inhibitor of glycogen synthase kinase-3 (GSK-3). GSK-3 inhibitors have been shown to be effective in reducing the lysosomal pH. See Avrahami et al., 2013, *Commun Integr Biol* 6(5): e25179, which is herein incorporated by reference in its entirety. For instance, the competitive GSK-3 inhibitor, L803-mts, has been shown to facilitate acidification of the lysosome by inhibiting GSK-3 activity, which acts to impair lysosomal acidification. Accordingly, in one embodiment, the inhibitor of GSK-3 is the cell permeable peptide, L803-mts (SEQ ID NO: 72). Suitable GSK-3 inhibitors may be found in US Patent Publication Nos. 2013/0303441 and 2015/0004255, which are herein incorporated by reference in their entireties. In some embodiments, the GSK-3 inhibitor is selected from 2'Z,3'E)-6-bromoindirubin-3'-acetoxime, TDZD-8 (4-Benzyl-2-methyl-1,2,4-thiadiazolidine-3,5-dione), SB216763 (3-(2,4-Dichlorophenyl)-4-(1-methyl-1H-indol-3-yl), NP-103, 2-Thio(3-iodobenzyl)-5-(1-pyridyl)-[1,3,4]-oxadiazole, L803, L803-mts, and GF-109203X (2-[1-(3-Dimethylaminopropyl)indol-3-yl]-3-(indol-3-yl)malemide and pharmaceutically acceptable salts and mixtures thereof.

[0135] In some embodiments, the methods of the present invention further comprise the administration of one or more drugs that promotes autophagy. As used herein, drugs that promote autophagy can promote the intracellular degradation system that delivers cytoplasmic constituents to the lysosome. Accordingly, in some embodiments, the present invention provides a method of treating or preventing amyloidosis in a subject comprising administering to the subject a composition comprising a therapeutically effective amount

of at least one catabolic enzyme or a biologically active fragment thereof, and one or more drugs that promotes autophagy. In some embodiments, the present invention provides a method of treating or preventing amyloidosis in a subject comprising administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof, wherein the subject is also administered one or more drugs that acidifies the lysosome and/or endosome, and one or more drugs that promotes autophagy. In some embodiments, the drug that acidifies the lysosome and/or endosome, and the drug that promotes autophagy can be the same drug, or different drugs. As described herein, when performing a combination therapy, the drugs (e.g., a catabolic enzyme or a biologically active fragment thereof, a drug that acidifies the lysosome and/or endosome, and/or a drug that promotes autophagy) can be administered simultaneously or sequentially in any order. Without wishing to be bound by any particular theory, a treatment of therapeutic catabolic enzyme or a biologically active fragment thereof with an agent that can cause lysosome and/or endosome acidification and/or an agent that can promote autophagy is capable of lowering pH to optimal conditions for enzymatic proteolysis, and improving lysosomal proteolysis power.

[0136] In some embodiments, autophagy promoting reagents include, but are not limited to reagents that directly or indirectly promote autophagy such as TFEB activators, PPAR agonists, PGC-1α activators, LSD1 inhibitors, mTOR inhibitors, GSK3 inhibitors, etc.

[0137] In some embodiments, the drug promotes autophagy via activation of Transcription factor EB (TFEB) pathway. TFEB is a master gene for lysosomal biogenesis. It encodes a transcription factor that coordinates expression of lysosomal hydrolases, membrane proteins and genes involved in autophagy. TFEB overexpression in cultured cells induced lysosomal biogenesis and increased the degradation of complex molecules. TFEB is activated by PGC-1α and promotes reduction of htt aggregation and neurotoxicity.

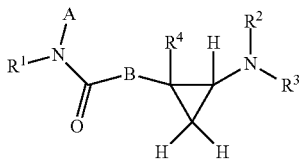
[0138] In some embodiments, the drug that promotes autophagy via activation of TFEB pathway is an activator of TFEB. In some embodiments, such TFEB activator include, but are not limited to C1 (Song et al, 2016, Autophagy, 12(8):1372-1389), and 2-hydroxypropyl-β-cyclodextrin (Kilpatrick et al., 2015, PLOS ONE DOI:10.1371/journal.pone.0120819). Each of the references mentioned herein is incorporated by reference in its entirety.

[0139] In some embodiments, the drug that promotes autophagy via activation of TFEB pathway is an agent that can activate peroxisome proliferator-activated receptor gamma coactivator 1-α (PGC-1α). In some embodiments, such activators of PGC-1α include, but are not limited to, pyrroloquinoline quinone, resveratrol, R-α-lipoic acid (ALA), ALA /acetyl-L-carnitine (ALC), flavonoids, isoflavones and derivatives (e.g., quercetin, daidzein, genistein, biochanin A, and formononetin). See, Das and Sharma 2015 (CNS & Neurological Disorders—Drug Targets, 2015, 14, 1024-1030.) Each of the references mentioned herein is incorporated by reference in its entirety.

[0140] In some embodiments, the drug promotes autophagy via activation of peroxisome proliferator-activated receptor gamma coactivator 1-α (PGC-1α) and/or Forehead box O3 (FOXO3). PGC-1α is a master regulator of mitochondrial biogenesis. PGC-1α interacts with the nuclear receptor PPAR-γ, which permits the interaction of this protein with multiple transcription factors. This protein can interact with, and regulate the activities of, cAMP response element-binding protein (CREB) and nuclear respiratory factors (NRFs). It provides a direct link between external physiological stimuli and the regulation of mitochondrial biogenesis, and is a major factor that regulates muscle fiber type determination. FOXO3 is a transcription factor that can be inhibited and translocated out of the nucleus on phosphorylation by protein such as Akt/PKB in the PI3K signaling pathway.

[0141] In some embodiments, a drug that promotes autophagy via PGC-1α and/or FOXO3 activation is an inhibitor of Lysine (K)-specific demethylase 1A (LSD1). LSD1 is a flavin-dependent monoamine oxidase, which can demethylate mono- and bi-methylated lysines. LSD1 has roles critical in embryogenesis and tissue-specific differentiation. In some embodiments, such LSD1 inhibitors include, but are not limited to, 1-(4-methyl-1-piperazinyl)-2-[[(1R*,2S*)-2-[4-phenylmethoxy)phenyl]cyclopropyl] amino]ethanone dihydrochloride (RN-1; Cui et al., 2015, Blood 2015 126:386-396), CBB1001-1009 (Wang et al., 2011, Cancer Res. 2011 Dec. 1; 71(23): 7238-7249.), TCP, Pargyline, CGC-11047, and Namolone (Pieroni et al., 2015, European Journal of Medicinal Chemistry 92 (2015) 377e386), phenelzine analogues (Prusevich et al., ACS Chem. Biol. 2014, 9, 1284-1293), and those described in WO2015156417, which is herein incorporated by reference in its entirety. In some embodiments, one or more LSD1 inhibitors are used. In some embodiments, both RN-1 and a LSD1 inhibitor described in WO2015156417 are used. WO2015156417 describes inhibitors of LSD1 represented by formula I:

(I)

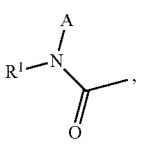


wherein, A is an optionally substituted heterocyclic group, or an optionally substituted hydrocarbon group; B is a ring selected from

[0142] (1) a 5- or 6-membered aromatic heterocycle optionally fused with an optionally substituted 5- or 6-membered ring, and

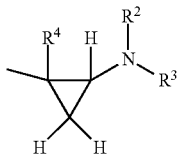
[0143] (2) a benzene ring fused with an optionally substituted 5- or 6-membered ring, wherein the ring represented by B is optionally substituted, and binds, via two adjacent carbon atoms with one atom in between, to a group represented by the formula

(II)



and a group represented by the formula

(III)



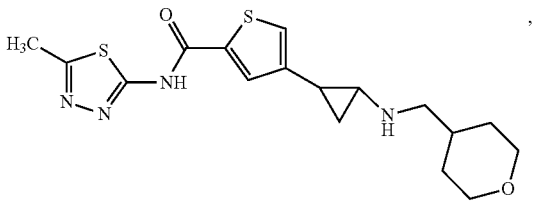
[0144] R^1, R^2, R^3 and R^4 are each independently a hydrogen atom, an optionally substituted hydrocarbon group or an optionally substituted heterocyclic group;

[0145] A and R^1 are optionally bonded with each other to form, together with the adjacent nitrogen atom, an optionally substituted cyclic group; and

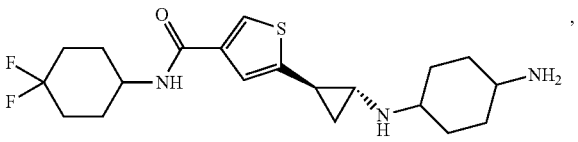
[0146] R^2 and R^3 are optionally bonded with each other to form, together with the adjacent nitrogen atom, an optionally substituted cyclic group, or a salt thereof. Such LSD1 inhibitors are more specific with less side effect and good blood-brain barrier penetration.

[0147] In some embodiments, the LSD1 inhibitors are selected from the group consisting of the following compounds (compounds 1-30), and salts, stereoisomers, geometric isomers, tautomers, oxynitrides, enantiomers, diastereoisomers, racemates, prodrugs, solvates, metabolites, esters, and mixtures thereof:

Compound 1

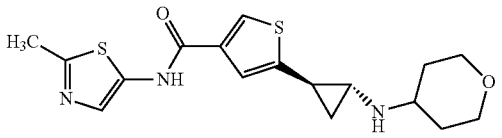


Compound 2

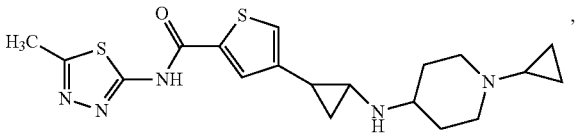


(cyclopropane: (1R,2R) cyclohexane: cis or trans)

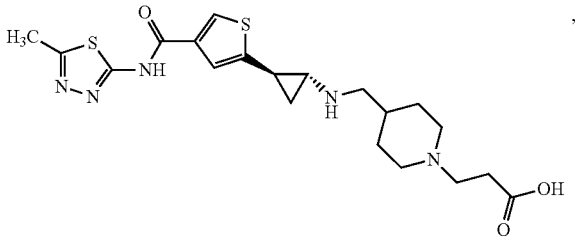
Compound 3



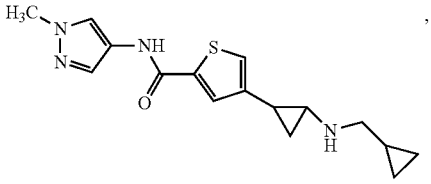
Compound 4



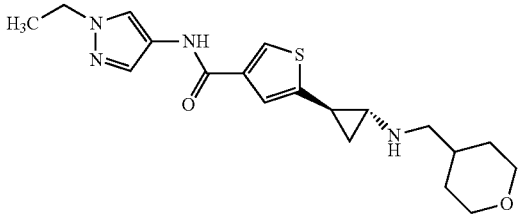
Compound 5



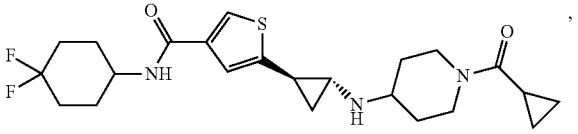
Compound 6



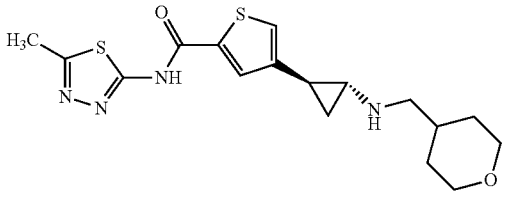
Compound 7



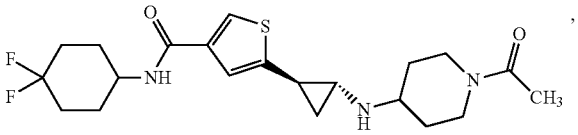
Compound 8



Compound 9

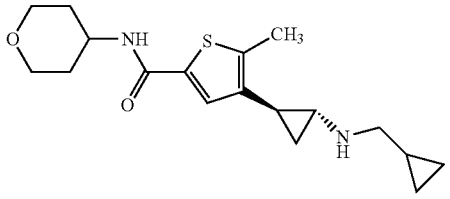


Compound 10

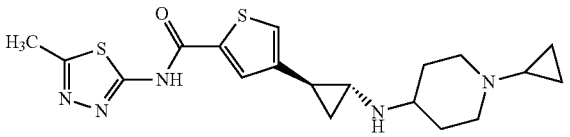


-continued

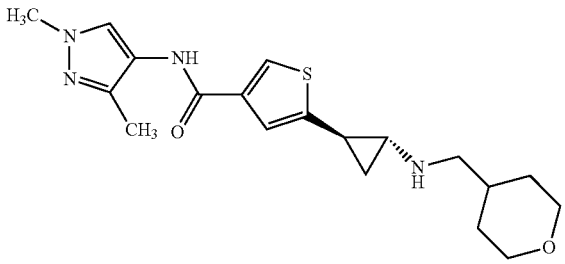
Compound 11



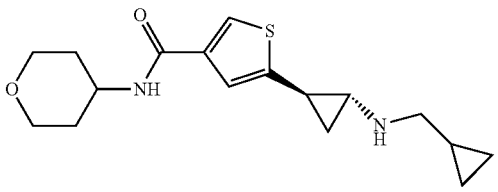
Compound 12



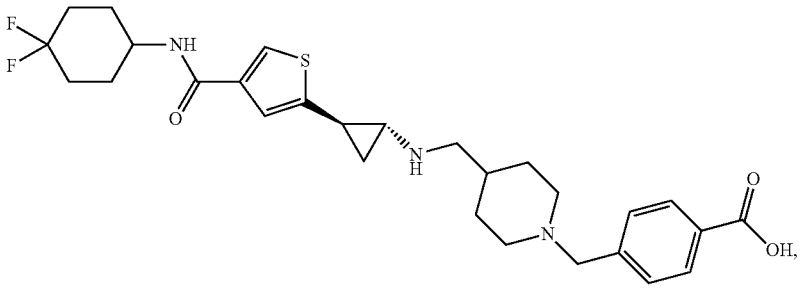
Compound 13



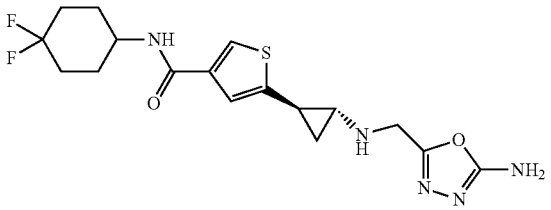
Compound 14



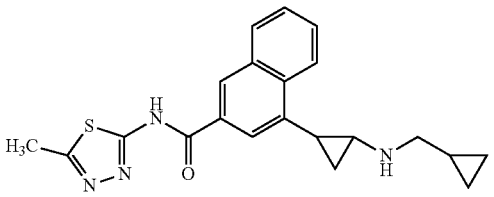
Compound 15



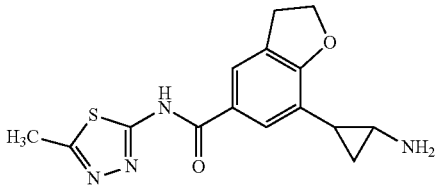
Compound 16



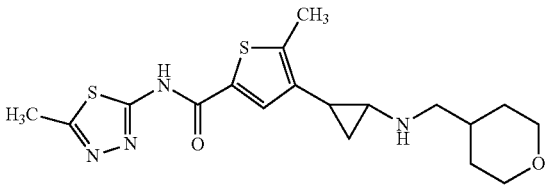
Compound 17



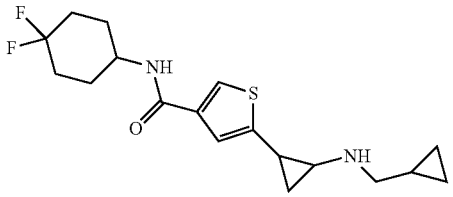
Compound 18



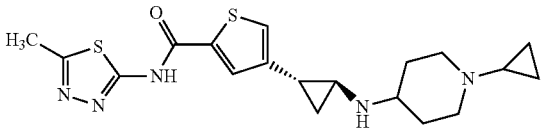
Compound 19



Compound 20

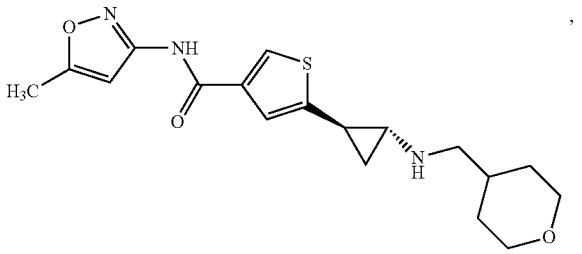


Compound 21

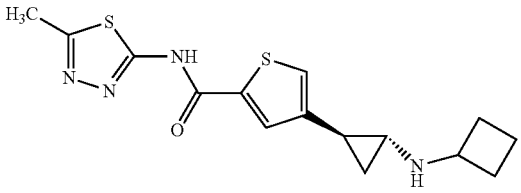


-continued

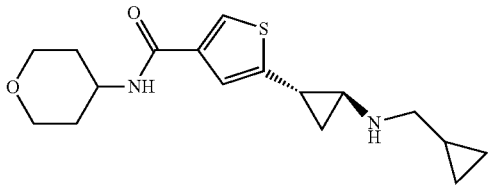
Compound 22



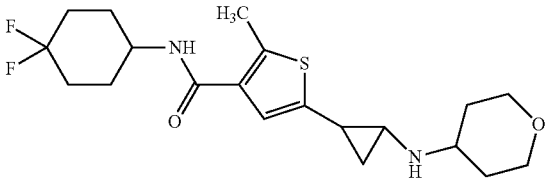
Compound 23



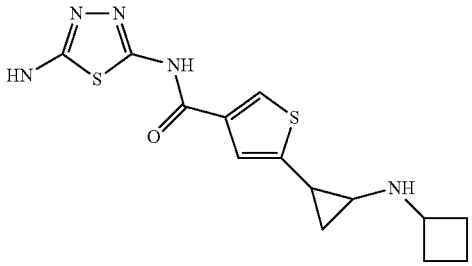
Compound 24



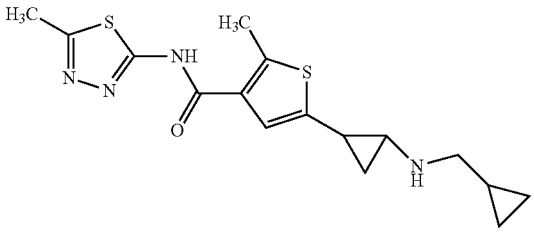
Compound 25



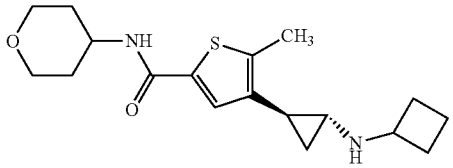
Compound 26



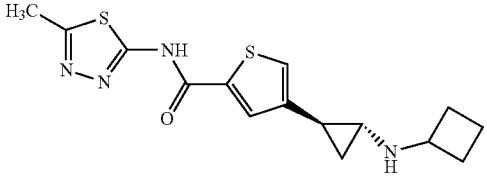
Compound 27



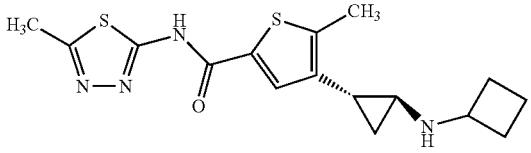
Compound 28



Compound 29



Compound 30



In one embodiment, the LSD1 inhibitor to be co-administered with a catabolic enzyme of the present invention or a biologically active fragment thereof is compound 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, or any mixtures thereof.

[0148] In some embodiments, the drug is capable of modify the activity of a regulator or a co-activator of PGC-1α. Such regulators or co-activators of PGC-1α include, but are not limited to, Parkin Interacting Substrate (PARIS), Sirtuin 1 (SIRT1), 5' AMP-activated protein kinase (AMPK), General control of amino acid synthesis protein 5 (GCN5), Nuclear respiratory factor 1, 2(NRF-1,2), Glycogen synthase kinase 3β (GSK3β), Peroxisome proliferator-activated receptor-α,β/δ,γ (PPAR-α,β/δ,γ), p38 mitogen-activated protein kinase (p38MAPK), Estrogen-related receptors (ERRs), myocyte enhancer factor-2 (MEF2), and Thyroid hormone receptor (TR), see Das and Sharma (CNS & Neurological Disorders—Drug Targets, 2015, 14, 1024-1030). Each of the references mentioned herein is incorporated by reference in its entirety.

[0149] In some embodiments, the drug that promotes autophagy is a Peroxisome proliferator-activated receptor (PPAR) agonist. PPARs are nuclear receptor proteins that function as transcription factors regulating the expression of genes. They are critical in the regulation of cellular differentiation, development, and metabolism and tumorigenesis.

[0150] In some embodiments, the PPAR is selected from PPARα, PPARβ/δ, and PPARγ. In some embodiments, the PPAR agonist is a PPARα agonist, including but not limited to amphipathic carboxylic acids (e.g., clofibrate, gemfibro-

zil, ciprofibrate, bezafibrate, and fenofibrate), fibrate, ureidofibrate, oxybenzylglycine, triazolone, agonists containing a 2,4-dihydo-3H-1,2,4 triazole-3-one (triazolone) core (e.g., LY518674), BMS-687453, Wy-14643, GW2331, GW 95798, LY518674, and GW590735.

[0151] In some embodiments, the PPAR agonist is a PPARβ/δ agonist, including but not limited to GW501516 (Brunmair; et al. Diabetologia. 49 (11): 2713-22), L-165041, compound 7 (Burdick et al., Cell Signal 2006, 18 (1), 9-20), thiazole, bisaryl substituted thiazoles, non-TZD compounds (e.g., L-165041), L-165041, compound 7 (Burdick et al., Cell Signal 2006, 18 (1), 9-20), 38c (Johnson et al., J Steroid Biochem Mol Biol 1997, 63 (1-3), 1-8), and oxazoles. Each of the references mentioned herein is incorporated by reference in its entirety.

[0152] In some embodiments, the PPAR agonist is a PPARγ agonist, including but not limited to thiazolidinediones (TZDs or glitazones), glitazar, indenone, NSAIDs, dihydrocinnamate, β-carboxyethyl rhodamine, and those described in Corona and Duchen, 2016 (Free Radical Biology and Medicine, published online Jun. 23, 2016). In some embodiments, the PPARγ agonist is an endogenous or natural agonist. In some embodiments, the PPARγ agonist is a synthetic agonist. In some embodiments, the PPARγ agonist is selected from the group consisting of eicosanoids prostaglandin-A1, cyclopentenone prostaglandin 15-deoxy-$\Delta^{12,14}$-Prostaglandin J2 (15D-PGJ2), unsaturated fatty acids such as linoleic acid and socosahexaenoic acid, nitroalkenes such as nitrated oleic acid and linoleic acid, oxidized phospholipids such as hexadecyl azelaoyl phosphatidylcholine and lysophosphatidic acid, non-steroidal anti-inflammatory drugs, such as flufenamic acid, ibuprofen, fenoprofen, and indomethacin, pioglitazone, GW0072, ciglitazone, troglitazone, rosiglitazone, isoglitazone, NC-2100 (Loiodice et al., Curr. Top. Med. Chem. 2011, 11(7):819-39), SB-236636, tesaglitazar, farglitazar, GW1929, compound 14c (Haigh et al., Bioorg Med Chem 1999, 7(5):821-30), SP1818, ragaglitazar, metaglidasen, balaglitazone, and INT131. Each of the references mentioned herein is incorporated by reference in its entirety.

[0153] In some embodiments, the PPAR agonist binds to PPARα, PPARβ/δ, and PPARγ, such as bezafibrate, LY465608, indeglitazar, TIPP-204, GW693085, TIPP-401, and TIPP-703. In some embodiments, the PPAR agonist binds to PPARα and PPARγ, such as farglitazar, muraglitazar, tesaglitazar, GW409544, aleglitazar, MK-767, TAK-559, compound 18 (Kojo et al., J. Pharmacol Sci 2003, 93 (3), 347-55), compounds 68, 70, 72, 76 (Felts et al., J Med Chem 2008, 51 (16), 4911-9), metaglidasen, and S-2/S-4 (Suh et al., J Med Chem 2008, 51 (20), 6318-33). In some embodiments, the PPAR agonist binds to PPARβ and PPARγ, such as compound 23 (Martin et al., J Med Chem 2009, 52(21), 6835-50). More PPARs agonists are described in Nevin et al., 2011 (Current Medicinal Chemistry, 2011, 18, 5598-5623). Each of the references mentioned herein is incorporated by reference in its entirety.

[0154] In some embodiments, the drug that promotes autophagy is an inhibitor of mechanistic target of rapamycin (mTOR). mTOR is a serine/threonine-specific protein kinase that belongs to the family of phosphatidylinositol-3 kinase (PI3K) related kinases (PIKKs), see Maiese et al. (Br J Clin Pharmacol, 82(5):1245-1266), which is herein incorporated by reference in its entirety. mTOR integrates the input from upstream pathways, including insulin, growth factors (such as IGF-1 and IGF-2), and amino acids, and also senses cellular nutrient, oxygen, and energy levels. In some embodiments, mTOR inhibitors include, but are not limited to, an antibody of mTOR, rapamycin and its analogs (e.g., temsirolimus (CCI-779), everolimus (RAD001), ridaforolimus (AP-23573), sirolimus, deforolimus), curcumin (Zhang et al., 2016, Oncotarget), curcumin analogs (Song et al. 2016, Autophagy, 12(8):1372-1389), ATP-competitive mTOR kinase inhibitors, mTOR/PI3K dual inhibitors (dactolisib, BGT226, SF1126, PKI-587 etc.), deptor (Maiese, Neural Regeneration Research. 2016; 11(3):372-385), and mTORC1/mTORC2 dual inhibitors (TORCdIs, such as sapanisertib (a.k.a. INK128), AZD8055, and AZD2014). Each of the references mentioned herein is incorporated by reference in its entirety.

[0155] In some embodiments, the drug that promotes autophagy is an inhibitor of Glycogen synthase kinase 3 (GSK3). GSK3 is a serine/threonine protein kinase that mediates the addition of phosphate molecules onto serine and threonine amino acid residues. In some embodiments, the GSK3 inhibitor is ATP-competitive. In some embodiments, the GSK3 inhibitor is non-ATP competitive. In some embodiments, GSK3 inhibitors include, but are not limited to, an antibody of GSK3, metal cations (e.g., beryllium, copper, lithium, mercury, and tungsten), marine organism-derived drugs (e.g., 6-BIO, dibromocantharelline, hymenialdesine, indirubins, meridianins, manzamine A, palinurine, tricantine), aminopyrimidines (e.g., CT98014, CT98023, CT99021, and TWS119), ketamine, arylindolemaleimide (e.g., SB-216763 and SB-41528), thiazoles (e.g., AR-A014418 and AZD-1080), paullones (e.g., Alsterpaullone, Cazpaullone, Kenpaullone), thiadiazolidindiones (e.g., TDZD-8, NP00111, NP031115, and tideglusib), halomethylketones (e.g., HMK-32), certain peptides (L803-mts), SB415286, SB216763, and CT99021 (Stretton et al., 2015, Biochem. J. (2015) 470, 207-221; Marchand et al., 2015, The Journal of Biological Chemistry, 290(9):5592-5605). Each of the references mentioned herein is incorporated by reference in its entirety.

[0156] In some embodiments, the methods of the present invention further comprise the administration of one or more drugs that modulates the lysosome. In some embodiments, drugs that modulate the lysosome may be capable of decreasing the level of Rab5a, a marker of early endosomes. Accordingly, in some embodiments, the present invention provides a method of treating or preventing amyloidosis in a subject comprising administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof, wherein the subject is also administered one or more drugs that modulates the lysosome. As described herein, when performing a combination therapy, the two or more drugs (e.g., a catabolic enzyme or a biologically active fragment thereof and a drug that modulates the lysosome) can be administered simultaneously or sequentially in any order

[0157] In some embodiments, the drug that modulates the lysosome is Z-phenylalanyl-alanyl-diazomethylketone (PADK) or a PADK analog, or a pharmaceutically acceptable salt or ester thereof. In some embodiments, the PADK analog is selected from Z-L-phenylalanyl-D-alanyl-diazomethylketone (PdADK), Z-D-phenylalanyl-L-alanyl-diazomethylketone (dPADK), and Z-D-phenylalanyl-D-alanyl-diazomethylketone (dPdADK). In some embodiments, the

drug that modulates the lysosome is Z-phenylalanyl-phenylalanyl-diazomethylketone (PPDK) or a PPDK analog, or a pharmaceutically acceptable salt or ester thereof. An exemplary listing of suitable lysosome modulators may be found in US Patent Publication No. 2016/0136229, which is herein incorporated by reference in its entirety.

[0158] In some embodiments, when performing a combination therapy, the two or more drugs can be administered simultaneously or sequentially in any order. In some embodiments, when at least two drugs are administered sequentially, the duration between the two administrations can be about 1 minute, 5 minutes, 10 minutes, 20 minutes, 30 minutes, 1 hour, 2 hours, 4 hours, 6 hours, 12 hours, 24 hours, 2 days, three days, 1 week, 2 weeks, 3 weeks, 1 month, 2 months, 3 months, or more.

[0159] In some embodiments, the methods of the present invention further comprise a surgery to be performed on the subject. In some embodiments, the surgery is stem cell transplantation and/or organ transplantation. In some embodiments, the stem cell transplantation is autologous (e.g., stem cells derived from the subject).

[0160] In some embodiments, the methods further comprise providing a supportive treatment to the subject. In some embodiments, when the heart or kidneys of the subject are affected, the methods comprise taking a diuretic (water excretion pill), restricting the amount of salt in diet, and/or wearing elastic stockings and elevating their legs to help lessen the amount of swelling. In some embodiments, when the gastrointestinal tract is involved, dietary changes and certain medications can be tried to help symptoms of diarrhea and stomach fullness.

[0161] A pharmaceutical composition of the present invention can be administered to a patient by any suitable methods known in the art. In some embodiments, administration of a composition of the present invention may be carried out orally, parenterally, subcutaneously, intravenously, intramuscularly, intraperitoneally, by intranasal instillation, by implantation, by intracavitary or intravesical instillation, intraocularly, intraarterially, intralesionally, transdermally, aerosolly (e.g., inhalation) or by application to mucous membranes.

[0162] In some embodiments, a pharmaceutical composition of the present invention further comprises a pharmaceutically-acceptable carrier. When the term "pharmaceutically acceptable" is used to refer to a pharmaceutical carrier or excipient, it is implied that the carrier or excipient has met the required standards of toxicological and manufacturing testing or that it is included on the Inactive Ingredient Guide prepared by the U.S. Food and Drug administration.

[0163] Compositions intended for oral use may be prepared in either solid or fluid unit dosage forms. Fluid unit dosage form can be prepared according to procedures known in the art for the manufacture of pharmaceutical compositions and such compositions may contain one or more agents selected from the group consisting of sweetening agents, flavoring agents, coloring agents and preserving agents in order to provide pharmaceutically elegant and palatable preparations. An elixir is prepared by using a hydroalcoholic (e.g., ethanol) vehicle with suitable sweeteners such as sugar and saccharin, together with an aromatic flavoring agent. Suspensions can be prepared with an aqueous vehicle with the aid of a suspending agent such as acacia, tragacanth, methylcellulose and the like.

[0164] Solid formulations such as tablets contain the active ingredient in admixture with non-toxic pharmaceutically acceptable excipients that are suitable for the manufacture of tablets. These excipients may be for example, inert diluents, such as calcium carbonate, sodium carbonate, lactose, calcium phosphate or sodium phosphate: granulating and disintegrating agents for example, corn starch, or alginic acid: binding agents, for example starch, gelatin or acacia, and lubricating agents, for example magnesium stearate, stearic acid or talc and other conventional ingredients such as dicalcium phosphate, magnesium aluminum silicate, calcium sulfate, starch, lactose, methylcellulose, and functionally similar materials. The tablets may be uncoated or they may be coated by known techniques to delay disintegration and absorption in the gastrointestinal tract and thereby provide a sustained action over a longer period. For example, a time delay material such as glyceryl monostearate or glyceryl distearate may be employed.

[0165] Formulations for oral use may also be presented as hard gelatin capsules wherein the active ingredient is mixed with an inert solid diluent, for example, calcium carbonate, calcium phosphate or kaolin, or as soft gelatin capsules wherein the active ingredient is mixed with water or an oil medium, for example peanut oil, liquid paraffin or olive oil. Soft gelatin capsules are prepared by machine encapsulation of a slurry of the compound with an acceptable vegetable oil, light liquid petrolatum or other inert oil.

[0166] Aqueous suspensions contain active materials in admixture with excipients suitable for the manufacture of aqueous suspensions. Such excipients are suspending agents, for example sodium carboxylmethylcellulose, methyl cellulose, hydropropylmethylcellulose, sodium alginate, polyvinylpyrrolidone, gum tragacanth and gum acacia: dispersing or wetting agents may be a naturally-occurring phosphatide, for example, lecithin, or condensation products of an alkylene oxide with fatty acids, for example polyoxyethylene stearate, or condensation products of ethylene oxide with long chain aliphatic alcohols, for example hepta-decaethyleneoxycetanol, or condensation products of ethylene oxide with partial esters derived from fatty acids and a hexitol such as polyoxyethylene sorbitol monooleate, or condensation products of ethylene oxide with partial esters derived from fatty acids and hexitol anhydrides, for example polyethylene sorbitan monooleate. The aqueous suspensions may also contain one or more preservatives, for example ethyl, or n-propyl-p-hydroxy benzoate, one or more colouring agents, one or more flavoring agents or one or more sweetening agents, such as sucrose or saccharin.

[0167] Oily suspensions may be formulated by suspending the active ingredients in a vegetable oil, for example peanut oil, olive oil, sesame oil or coconut oil, or in a mineral oil such as liquid paraffin. The oily suspensions may contain a thickening agent, for example beeswax, hard paraffin or cetyl alcohol. Sweetening agents such as those set forth above, and flavoring agents may be added to provide palatable oral preparations. These compositions may be preserved by the addition of an anti-oxidant such as ascorbic acid.

[0168] Dispersible powders and granules suitable for preparation of an aqueous suspension by the addition of water provide the active ingredient in admixture with a dispersing or wetting agent, suspending agent and one or more preservatives. Suitable dispersing or wetting agents and suspending agents are exemplified by those already

mentioned above. Additional excipients, for example sweetening, flavoring and colouring agents, may also be present.

[0169] Pharmaceutical compositions of the invention may also be in the form of oil-in-water emulsions. The oil phase may be a vegetable oil, for example olive oil or peanut oil, or a mineral oil, for example liquid paraffin or mixtures of these. Suitable emulsifying agents may be naturally-occurring gums, for example gum acacia or gum tragacanth, naturally-occurring phosphatides, for example soy bean, lecithin, and esters or partial esters derived from fatty acids and hexitol, anhydrides, for example sorbitan monooleate, and condensation products of the said partial esters with ethylene oxide, for example polyoxyethylene sorbitan monooleate. The emulsions may also contain sweetening and flavoring agents.

[0170] The pharmaceutical compositions may be in the form of a sterile injectable aqueous or oleaginous suspension. This suspension may be formulated according to known art using those suitable dispersing or wetting agents and suspending agents that have been mentioned above. The sterile injectable preparation may also be a sterile injectable solution or a suspension in a non-toxic parentally acceptable diluent or solvent, for example as a solution in 1,3-butanediol. Among the acceptable vehicles and solvents that may be employed are water, Ringer's solution and isotonic sodium chloride solution. In addition, sterile, fixed oils are conventionally employed as a solvent or suspending medium. For this purpose any bland fixed oil may be employed including synthetic mono- or diglycerides. In addition, fatty acids such as oleic acid find use in the preparation of injectables. Adjuvants such as local anaesthetics, preservatives and buffering agents can also be included in the injectable solution or suspension.

[0171] In some embodiments, the delivery systems suitable include time-release, delayed release, sustained release, or controlled release delivery systems. In some embodiments, a composition of the present invention can be delivered in a controlled release system, such as sustained-release matrices. Non-limiting examples of sustained-release matrices include polyesters, hydrogels (e.g., poly(2-hydroxyethyl-methacrylate) as described by Langer et al., 1981, J. Biomed. Mater. Res., 15:167-277 and Langer, 1982, Chem. Tech., 12:98-105), or poly(vinylalcohol)], polylactides (U.S. Pat. No. 3,773,919; EP 58,481), copolymers of L-glutamic acid and gamma ethyl-L-glutamate (Sidman et al., 1983, Biopolymers, 22:547-556), non-degradable ethylene-vinyl acetate (Langer et al., supra), degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT™ (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(−)-3-hydroxybutyric acid (EP 133,988). In some embodiments, the composition may be administered using intravenous infusion, an implantable osmotic pump, a transdermal patch, liposomes, or other modes of administration. In one embodiment, a pump may be used (see Langer, supra; Sefton, CRC Crit. Ref. Biomed. Eng. 14:201 (1987); Buchwald et al., Surgery 88:507 (1980); Saudek et al., N. Engl. J. Med. 321:574 (1989). In another embodiment, polymeric materials can be used. In yet another embodiment, a controlled release system can be placed in proximity to the therapeutic target, for example liver, thus requiring only a fraction of the systemic dose (see, e.g., Goodson, in Medical Applications of Controlled Release, supra, vol. 2, pp. 115-138 (1984). Other controlled release systems are discussed in the review by Langer (Science 249:1527-1533 (1990). In some embodiments, the composition may be administered through subcutaneous injection.

[0172] In some embodiments, the release of the composition occurs in bursts. Examples of systems in which release occurs in bursts includes, e.g., systems in which the composition is entrapped in liposomes which are encapsulated in a polymer matrix, the liposomes being sensitive to specific stimuli, e.g., temperature, pH, light or a degrading enzyme and systems in which the composition is encapsulated by an ionically-coated microcapsule with a microcapsule core degrading enzyme.

[0173] In some embodiments, the release of the composition is gradual/continuous. Examples of systems in which release of the inhibitor is gradual and continuous include, e.g., erosional systems in which the composition is contained in a form within a matrix and effusional systems in which the composition is released at a controlled rate, e.g., through a polymer. Such sustained release systems can be e.g., in the form of pellets, or capsules.

[0174] Other embodiments of the compositions administered according to the invention incorporate particulate forms, protective coatings, protease inhibitors or permeation enhancers for various routes of administration, such as parenteral, pulmonary, nasal and oral. Other pharmaceutical compositions and methods of preparing pharmaceutical compositions are known in the art and are described, for example, in "*Remington: The Science and Practice of Pharmacy*" (formerly "*Remingtons Pharmaceutical Sciences*"); Gennaro, A., Lippincott, Williams & Wilkins, Philadelphia, Pa. (2000). In some embodiments, the pharmaceutical composition may further include a pharmaceutically acceptable diluent, excipient, carrier, or adjuvant.

[0175] In some embodiments, the dosage to be administered is not subject to defined limits, but it will usually be an effective amount, or a therapeutically/pharmaceutically effective amount. The term "effective amount" refers to the amount of one or more compounds that renders a desired treatment outcome. An effective amount may be comprised within one or more doses, i.e., a single dose or multiple doses may be required to achieve the desired treatment endpoint. The term "therapeutically/pharmaceutically effective amount" as used herein, refers to the level or amount of one or more agents needed to treat a condition, or reduce or prevent injury or damage, optionally without causing significant negative or adverse side effects. It will usually be the equivalent, on a molar basis of the pharmacologically active free form produced from a dosage formulation upon the metabolic release of the active free drug to achieve its desired pharmacological and physiological effects. In some embodiments, the compositions may be formulated in a unit dosage form. The term "unit dosage form" refers to physically discrete units suitable as unitary dosages for human subjects and other mammals, each unit containing a predetermined quantity of active material calculated to produce the desired therapeutic effect, in association with a suitable pharmaceutical excipient.

[0176] In some embodiments, dosing regimen of a pharmaceutical composition of the present invention includes, without any limitation, the amount per dose, frequency of dosing, e.g., per day, week, or month, total amount per dosing cycle, dosing interval, dosing variation, pattern or modification per dosing cycle, maximum accumulated dosing, or warm up dosing, or any combination thereof.

[0177] In some embodiments, dosing regimen includes a pre-determined or fixed amount per dose in combination with a frequency of such dose. For example, dosing regimen includes a fixed amount per dose in combination with the frequency of such dose being administered to a subject.

[0178] In some embodiments, the at least one catabolic enzyme (e.g., PPCA, NEU1, TPP1, cathepsin B, cathepsin D, cathepsin E, cathepsin K, and/or cathepsin L) is administered at about 0.1 to 20 mg/kg daily, weekly, biweekly, monthly, or bi-monthly. In some embodiments, the at least one intralysosomal catabolic enzyme is administered at about 0.2 to 15 mg/kg, about 0.5 to 12 mg/kg, about 1 to 10 mg/kg, about 2 to 8 mg/kg, or about 4 to 6 mg/kg daily, weekly, biweekly, monthly, or bi-monthly.

[0179] Based on the suitable dosage, the at least one catabolic enzyme can be provided in various suitable unit dosages. For example, a catabolic enzyme can comprise a unit dosage for administration of one or multiple times per day, for 1-7 days per week, or for 1-31 times per month. Such unit dosages can be provided as a set for daily, weekly and/or monthly administration.

[0180] As will be appreciated by those skilled in the art, the duration of the treatment methods depends on the type of amyloidosis being treated, any underlying diseases associated with amyloidosis, the age and conditions of the subject, how the subject responds to the treatment, etc.

[0181] In some embodiments, a person having risk of developing amyloidosis (e.g., a person who is genetically predisposed or previously had amyloidosis or associated diseases) can also receive prophylactic treatment of the present invention to inhibit or delay the development of amyloidosis and/or associated diseases.

[0182] The pharmaceutical composition of the present invention may also alleviate, reduce the severity of, or reduce the occurrence of, one or more of the symptoms associated with amyloidosis. In some embodiments, the symptoms are those associated with light-chain (AL) amyloidosis (primary systemic amyloidosis) and/or AA amyloidosis (secondary amyloidosis). In some embodiments, the symptoms include, but are not limited to, fluid retention, swelling, shortness of breath, fatigue, irregular heartbeat, numbness of hands and feet, rash, shortness of breath, swallowing difficulties, swollen arms or legs, esophageal reflux, constipation, nausea, abdominal pain, diarrhea, early satiety, stroke, gastrointestinal disorders, enlarged liver, diminished spleen function, diminished function of the adrenal and other endocrine glands, skin color change or growths, lung problems, bleeding and bruising problems, decreased urine output, diarrhea, hoarseness or changing voice, joint pain, and weakness. In some embodiments, the symptoms are those associated with amyloid-beta (Aβ) amyloidosis. In some embodiments, the symptoms include, but are not limited to, common symptoms of Alzheimer's disease, including memory loss, confusion, trouble understanding visual images and spatial relationships, and problems speaking or writing.

[0183] In some embodiments, the methods further comprise monitoring the response of the subject after administration to avoid severe and/or fatal immune-mediated adverse reactions due to over-dosage. In some embodiments, the administration of a pharmaceutical composition of the present invention is modified, such as reduced, paused or terminated if the patient shows persistent adverse reactions. In some embodiments, the dosage is modified if the patient fails to respond within about 1 day, 2 days, 3 days, 4 days, 5 days, 6 days, 1 week, 2 weeks or more from administration of first dose.

[0184] In some embodiments, a pharmaceutical composition of the present invention can ameliorate, treat, and/or prevent one or more conditions or associated symptoms described herein in a clinically relevant, statistically significant and/or persistent fashion. In some embodiments, administration of a pharmaceutical composition of the present invention provides statistically significant therapeutic effect for ameliorating, treating, and/or preventing one or more symptoms of amyloidosis. In one embodiment, the statistically significant therapeutic effect is determined based on one or more standards or criteria provided by one or more regulatory agencies in the United States, e.g., FDA or other countries. In some embodiments, the statistically significant therapeutic effect is determined based on results obtained from regulatory agency approved clinical trial set up and/or procedure.

[0185] In some embodiments, the statistically significant therapeutic effect is determined based on a patient population of at least 50, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1000, or more. In some embodiments, the statistically significant therapeutic effect is determined based on data obtained from randomized and double blinded clinical trial set up. In some embodiments, the statistically significant therapeutic effect is determined based on data with a p value of less than or equal to about 0.05, 0.04, 0.03, 0.02 or 0.01. In some embodiments, the statistically significant therapeutic effect is determined based on data with a confidence interval greater than or equal to 95%, 96%, 97%, 98% or 99%. In some embodiments, the statistically significant therapeutic effect is determined on approval of Phase III clinical trial of the methods provided by the present invention, e.g., by FDA in the US.

[0186] In some embodiment, the statistically significant therapeutic effect is determined by a randomized double blind clinical trial of a patient population of at least 50, 100, 200, 300 or 350; treated with a pharmaceutical composition of the present invention, but not in combination with any other agent. In some embodiment, the statistically significant therapeutic effect is determined by a randomized clinical trial of a patient population of at least 50, 100, 200, 300 or 350 and using any commonly accepted criteria for amyloidosis symptoms assessment.

[0187] In general, statistical analysis can include any suitable method permitted by a regulatory agency, e.g., FDA in the US or China or any other country. In some embodiments, statistical analysis includes non-stratified analysis, log-rank analysis, e.g., from Kaplan-Meier, Jacobson-Truax, Gulliken-Lord-Novick, Edwards-Nunnally, Hageman-Arrindel and Hierarchical Linear Modeling (HLM) and Cox regression analysis.

[0188] The invention also provides packaged pharmaceutical compositions or kits. In some embodiments, the packaged pharmaceutical compositions or kits include a therapeutically effective amount of an intralysosomal catabolic enzyme or a formulation comprising an intralysosomal catabolic enzyme of the present invention described herein. In some embodiments, the compound or formulation can increase the expression, activity, and/or concentration of at least one intralysosomal catabolic enzyme in a subject when the composition is administered to the subject. In some embodiments, the packaged pharmaceutical compositions or

kits further comprise in combination with a label or insert advising that the pharmaceutical compound or formulation be administered in combination with a second agent for treating or preventing amyloidosis described herein.

[0189] In some embodiments, the packaged pharmaceutical compositions or kits further comprise a therapeutically effective amount of a second agent described herein. In some embodiments, the packaged pharmaceutical compositions or kits is packaged in combination with a label or insert advising that the second agent be administered in combination with the intralysosomal catabolic enzyme or the formulation comprising an intralysosomal catabolic enzyme, or the compound or formulation that can increase the expression, activity, and/or concentration of at least one intralysosomal catabolic enzyme in a subject.

[0190] As used herein, the term "label or insert" includes, but is not limited to all written, electronic, or spoken communication with the subject, or with any person substantially responsible for the care of the subject, regarding the administration of the compositions of the present invention. An insert may further include information regarding co-administration of the compositions of the present invention with other compounds or compositions. Additionally, an insert may include instructions regarding administration of the compositions of the present invention before, during, or after a meal, or with/without food.

[0191] The following examples illustrate various aspects of the invention. The examples should, of course, be understood to be merely illustrative of only certain embodiments of the invention and not to constitute limitations upon the scope of the invention.

EXAMPLES

Example 1

Degradative Effects of Intralysosomal Catabolic Enzymes on Synthetic Amyloid Species

[0192] In this example, an in vitro study is performed to illustrate that intralysosomal enzymes such as PPCA (i.e., cathepsin A), cathepsin B, cathepsin D, and/or cocktail mixtures of two or more intralysosomal enzymes can be used for the treatment of amyloidosis. Without being bound by theory, it is hypothesized that delivery of PPCA, cathepsin B, cathepsin D, and other intralysosomal enzymes to lysosomes can assist in the degradation of abnormally accumulated amyloid species, e.g., Aβ-amyloid species before they can be transported into the extracellular space by exocytosis and be deposited as amyloid plaques.

[0193] This in vitro study shows the degradative effects of PPCA, cathepsin B, and cathepsin D on synthetic Aβ-amyloid species in a test tube.

[0194] First, in vitro aggregation assays of Aβ-amyloid species using synthetic Aβ-peptides is performed via a Thioflavin-T (THT) assay and western blot. FIG. **1** shows the aggregation of synthetic Aβ42 peptide and Aβ15-36 peptide (negative control) monitored by Thioflavin-T (THT) at physiological conditions (FIG. **1**A) or an acidic pH (FIG. **1**B). FIG. **2** shows the aggregation of Aβ42 amyloid species over time 24 hours as detected by western blot.

[0195] Second, prevention of the aggregation of synthetic Aβ-amyloid species by proteolytic degradation using PPCA, cathepsin B, and cathepsin D is tested via a Thioflavin-T (THT) assay and western blot. FIG. **3** shows that cathepsin A (i.e., PPCA) prevents the aggregation of Aβ42 amyloid. FIG. **4** shows that PPCA prevents the aggregation of Aβ42 amyloid in a dose dependent manner. FIG. **5** shows that PPCA prevents the aggregation of both high and low molecular weight species of Aβ42 amyloid. FIG. **6** shows that cathepsin B prevents the aggregation of Aβ42 amyloid. FIG. **7** shows that cathepsin B moderately prevents the aggregation of Aβ42 amyloid in a dose dependent manner. FIG. **8** shows that cathepsin B prevents the aggregation of low molecular weight species of Aβ42 amyloid and degrades Aβ42 monomers in a time-dependent manner. FIG. **9** shows that cathepsin B prevents the aggregation of Aβ42 amyloid.

[0196] Lastly, the ability of PPCA, cathepsin B, and cathepsin D to degrade pre-formed synthetic Aβ-amyloid species was tested. FIG. **10** shows that PPCA, cathepsin B, PPCA plus cathepsin B, and cathepsin D degrade high molecular weight oligomers/fibrils of Aβ42 amyloid. Cathepsin D degrades low molecular oligomers and completely eliminates Aβ42 monomers.

[0197] Example 1 Summary:

[0198] Experiments in Example 1 were designed to determine (1) whether the selected intralysosomal catabolic enzymes can prevent aggregation/formation of Aβ amyloid species (called prevention) and (2) whether the selected intralysosomal catabolic enzymes can degrade already preformed Aβ amyloid species (called degradation). Example 1 experiments have shown that Aβ42 amyloid species can be aggregated in vitro using synthetic Aβ42 peptides, and that this process can be monitored by THT assay (FIG. **1**) and/or western blot analysis (FIG. **2**). The THT assay allows for the monitoring of dynamic changes in Aβ42 aggregation upon treatment with degradative enzymes.

[0199] Data obtained from the experiments of Example 1 reveal that PPCA can efficiently prevent formation of Aβ42 amyloid species as shown by THT assay (FIG. **3**, FIG. **4**) and western blot (FIG. **5**), as well as degrade already pre-formed amyloid species (FIG. **10**). Prevention of amyloid formation and degradation by PPCA was efficient, reproducible and showed concentration dependent dynamics (FIG. **4**). Data obtained from experiments with cathepsin B showed moderate reduction in amyloid species formation as measured by THT (FIG. **6**). Western blot analysis revealed that cathepsin B prevents aggregation of low molecular weight Aβ42 species and degrades Aβ42 monomers in a time dependent manner (FIG. **8**). Experiments with the use of cathepsin D revealed strong prevention of aggregation of Aβ42 species, measured by THT (FIG. **9**). Cathepsin D also showed degradation of low molecular oligomers in pre-aggregated amyloid species and complete elimination Aβ42 monomers (FIG. **10**).

Example 2

Degradation of Aβ42 Oligomers and Fibrils by Cathepsin A, B, and D

[0200] In this example, two protocols specific for oligomer and fibril formation were applied to aggregate amyloid material to investigate which forms of Aβ42 species can be degraded by cathepsin A (PPCA), cathepsin B and cathepsin D. Aggregated oligomers and fibrils were then subjected to an enzymatic treatment followed by western blot analysis.

[0201] Initially, oligomers and fibrils were aggregated for a period of 7 days and material collected at different time points (days: 0, 1, 3 and 7) was subjected to SDS-PAGE electrophoresis followed by western blot analysis. In FIG. **11**, Aβ42 oligomers and Aβ42 fibrils were probed with oligomer specific antibody (A11), which does not recognize monomeric and fibril Aβ42 species. Various forms of oligomers were positively detected on western blot carrying material aggregated using both, oligomer formation and fibril formation protocols. A significant reduction in oligomer forms was observed at day 7 of fibril formation procedure (FIG. **11**, line 9), indicating a time dependent transition from oligomers to fibrils, undetectable by A11 antibody. In FIG. **12**, the same material as shown in FIG. **11** was probed with E610 antibody, which is specific for both oligomers and fibrils of Aβ42. A lack of fibrils at day 7 was observed when oligomer formation protocol was applied (FIG. **12**, line 4) and a strong appearance of fibrils at day 7 when fibril formation protocol was applied.

[0202] To study enzymatic degradation of oligomer species, Aβ42 oligomers were first aggregated for 9 days at pH 7.0 at 25° C. and then additionally incubated overnight at 37° C. in various pH, optimal for each of enzymes used in the study (pH 5.0 Cathepsin A, B and pH 3.5 Cathepsin D), with and without addition of enzymes. Western blot was probed with oligomer specific A11 antibody (FIG. **13**). Additional overnight aggregation of oligomers was observed at pH 5.0 as indicated by presence of higher molecular weight oligomers (lines 1, 2, 4, and 5) when compared to control line 9 (incubation for 9 days at 25° C.). In contrast, this aggregation was not observed for oligomers incubated overnight at pH 3.5. Overnight treatment of oligomers with 90 ng of cathepsin A at pH 5.0 and 37° C. resulted in degradation of the lowest oligomer band (line 4). Treatment of oligomers with 90 ng of cathepsin B and D did not reveal changes in intensity or size of oligomer band (lines 5, 6).

[0203] To study enzymatic degradation of fibril species, Aβ42 fibrils were first aggregated for 9 days at pH 7.0 at 25° C. and then additionally incubated overnight at 37 C in various pH, optimal for each of enzymes used in the study (pH 5.0 cathepsin A, B and pH 3.5 cathepsin D), with and without addition of enzymes. Western blot was probed with oligomer specific E610 antibody (FIG. **14**). Additional overnight aggregation of fibrils was observed in all pHs applied, as indicated by the presence of stronger/darker smear (lines 1, 2, 3) when compared to control line 9 (incubation for 9 days at 25° C.). Overnight treatment of fibrils with 90 ng of cathepsin A at pH 5.0 and 37° C. resulted in reduction/degradation of the fibril smear as well as degradation of oligomer species (line 4 compared to line 1). Overnight treatment of fibrils with 90 ng of cathepsin B at pH 5.0 and 37° C. resulted in weak reduction/degradation of the fibril smear (line 5 compared to line 2). Overnight treatment of fibrils with 90 ng of cathepsin D at pH 3.5 and 37° C. did not result in visible reduction/degradation of fibril smear or oligomer bands.

Example 3

Degradation of Aβ42 Monomers by Cathepsin A Monitored by ELISA

[0204] The purpose of this example is to assess whether cathepsin A can degrade Aβ42 peptides (monomers).

[0205] In this example, an enzymatic treatment of peptides with 90 ng of cathepsin A was carried out for 0-2 hr at 37° C. and pH 5.0. An identical experiment without the addition of cathepsin A was performed in parallel. In both cases, phenol red, an inhibitor of Aβ aggregation was used to prevent peptide aggregation into higher molecular weight species of amyloid. The effects of supplementation or lack of cathepsin A on Aβ42 monomers were measured using commercially available ELISA (SensoLyte® Anti-Human β-Amyloid (1-42) Quantitative ELISA, Colorimetric) at various time points (0, 10, 30, 60, 120 min). Sensolite ELISA consists of two antibodies: C-terminal capture antibody, which recognizes specifically human Aβ42 peptide but not Aβ40 or Aβ41 and N-terminal detection antibody. Because Cathepsin A is a carboxyl peptidase, Aβ42 monomers, if degraded, will be degraded from their C-terminus. This degradation would result in a lack of C-terminal amino acid 42 and in consequence lack of capture by C-terminus specific antibody, which should be visualized as a loos of fluorescent signal in ELISA. The ELISA read out for samples treated with cathepsin A revealed a loss of fluorescent signal already within first 10 min of treatment indicating degradation of Aβ42 monomers from the C-terminus by cathepsin A (FIG. **15**). Samples without supplementation of cathepsin A showed a strong fluorescent signal in ELISA indicating lack of C-terminal degradation in the absence of enzyme and thus efficient capture of Aβ42 monomers by C-terminus antibody.

Example 4

Degradation of Aβ40 Amyloid Species by Cath A

[0206] Aggregation experiments showed that Aβ40 amyloid species can be aggregated in vitro using synthetic Aβ40 peptides, and that this process can be monitored by THT assay (FIG. **16**). When compared with aggregation of Aβ42 peptides, Aβ40 showed much slower and less efficient rate of aggregation (FIG. **16**A).

[0207] Additional experiments were performed where THT assay was used to monitor dynamic changes in Aβ42 & Aβ40 aggregation upon treatment with degradative enzyme Cath A (FIG. **17**). Initial experiment aimed to measure the effect of Cath A treatment on aggregation of both Aβ42 & Aβ40 peptides in real time. To achieve this, Cath A was simultaneously incubated with corresponding peptides and THT reagent in separate reactions at conditions optimal for Cath A proteolysis. The above experiment revealed that in contrast to Aβ42 (FIG. **17**A), aggregation of Aβ40 amyloid is not affected by Cath A, in applied experimental settings, even when high concentration of enzyme is used (FIG. **17**B, C). Second experiment was carried out to investigate whether the result of the initial experiment is due to lack of proteolysis of Aβ40 by Cath A or whether the speed of such proteolysis is slower than the speed of Aβ40 aggregation and therefore no changes in THT fluorescence could be observed. In this experiment Aβ40 peptide was first incubated with Cath A for up to two hours in conditions optimal for Cath A proteolysis and followed by incubation with THT to measure aggregation. Obtained data revealed that Aβ40 peptide did not aggregate after pre-incubation with Cath A, proving its proteolysis (FIG. **18**).

[0208] To prove that observed loss of aggregation by Aβ40 peptide is caused by carboxypeptidase activity of Cath A, Aβ40 peptide was incubated for two hours at 37° C. at pH

5 with varying concentrations of Cath A. Subsequently, the reaction was transferred to an ELISA plate pre-coated with a C-terminal capture antibody, specifically for Aβ40 peptide only and was co-incubated with N-terminal detection antibody overnight at 4°. The results have shown progressively reduced binding of Aβ40 peptide to C-terminal capture antibody with increasing concentration of Cath A (FIG. **19**). This proves that C-terminus of Aβ40 peptide was removed by caboxyterminal activity of Cath A.

[0209] Aggregation of Aβ40 peptide into amyloid species was also monitored using Western Blot technique (FIG. **20**A). We were able to aggregate Aβ40 into high molecular weight fibrils but not oligomeric forms using aggregation process taking up to 9 days. An experiment was carried out in which Aβ40 was simultaneously incubated Cath A for up to 9 days during the process of fibril formation. Obtained results revealed that Cath A significantly prevents formation of high molecular weight fibrils due to its proteolytic action on Aβ40 amyloid (FIG. **20**B). Reduction of levels of monomeric Aβ40 form was also observed in this experiment (FIG. **20**C).

[0210] Unless defined otherwise, all technical and scientific terms herein have the same meaning as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any methods and materials, similar or equivalent to those described herein, can be used in the practice or testing of the present invention, the preferred methods and materials are described herein. All publications, patents, and patent publications cited are incorporated by reference herein in their entirety for all purposes.

[0211] The publications discussed herein are provided solely for their disclosure prior to the filing date of the present application. Nothing herein is to be construed as an admission that the present invention is not entitled to antedate such publication by virtue of prior invention.

[0212] While the invention has been described in connection with specific embodiments thereof, it will be understood that it is capable of further modifications and the application is intended to cover any variations, uses, or adaptations of the invention following, in general, the principles of the invention and including such departures from the disclosure as come within known or customary practice within the art to which the invention pertains and as may be applied to the essential features set forth and as follows in the scope of the appended claims.

SEQUENCE LISTING

SEQ ID NO: 1 Human PPCA mRNA, variant 1 mRNA

```
  1    agagtgcacc cgaatccacg ggctcggagg cagcagccat ctctcggcca tagggcaggc
 61    cagctggcgc cgggggctat tttgggcggc gggcaatgat ggtgaccgca aggcgacctt
121    gtaaggcatt tccccctga ctcccttccc cgagcctctg cccgggggtc ctagcgccgc
181    tttctcagcc atcccgccta caacttagcc gtccacaaca ggatcatctg atcgcgtgcg
241    cccgggctac gatctgcgag gcccgcggac cttgacccgg cattgaccgc caccgccccc
301    caggtccgta gggaccaaag aaggggcggg aggaagactg tcacgtggcg ccggagttca
361    cgtgactcgt acacatgact tccagtcccc gggcgcctcc tggagagcaa ggacgcgggg
421    gagcagagat gatccgagcc gcgccgccgc cgctgttcct gctgctgctg ctgctgctgc
481    tgctagtgtc ctgggcgtcc cgaggcgagg cagcccccga ccaggacgag atccagcgcc
541    tccccgggct ggccaagcag ccgtctttcc gccagtactc cggctacctc aaaggctccg
601    gctccaagca cctccactac tggtttgtgg agtcccagaa ggatcccgag aacagccctg
661    tggtgctttg gctcaatggg ggtcccggct gcagctcact agatgggctc ctcacagagc
721    atggcccctt cctggtccag ccagatggtg tcaccctgga gtacaacccc tattcttgga
781    atctgattgc caatgtgtta tacctggagt ccccagctgg ggtgggcttc tcctactccg
841    atgacaagtt ttatgcaact aatgacactg aggtcgccca gagcaatttt gaggcccttc
901    aagatttctt ccgcctcttt ccggagtaca agaacaacaa acttttcctg accggggaga
961    gctatgctgg catctacatc cccaccctgg ccgtgctggt catgcaggat cccagcatga
1021   accttcaggg gctggctgtg ggcaatggac tctcctccta tgagcagaat gacaactccc
1081   tggtctactt tgcctactac catggccttc tggggaacag gctttggtct tctctccaga
1141   cccactgctg ctctcaaaac aagtgtaact tctatgacaa caaagacctg gaatgcgtga
1201   ccaatcttca ggaagtggcc cgcatcgtgg gcaactctgg cctcaacatc tacaatctct
1261   atgccccgtg tgctggaggg gtgcccagcc attttaggta tgagaaggac actgttgtgg
1321   tccaggattt gggcaacatc ttcactcgcc tgccactcaa gcggatgtgg catcaggcac
1381   tgctgcgctc aggggataaa gtgcgcatgg acccccccctg caccaacaca acagctgctt
1441   ccacctacct caacaacccg tacgtgcgga aggccctcaa catcccggag cagctgccac
1501   aatgggacat gtgcaacttt ctggtaaact tacagtaccg ccgtctctac cgaagcatga
1561   actcccagta tctgaagctg cttagctcac agaaatacca gatcctatta tataatggag
1621   atgtagacat ggcctgcaat ttcatggggg atgagtggtt tgtggattcc ctcaaccaga
1681   agatggaggt gcagcgccgg ccctggttag tgaagtacgg ggacagcggg gagcagattg
1741   ccggcttcgt gaaggagttc tcccacatcg cctttctcac gatcaagggc gccggccaca
1801   tggttcccac cgacaagccc ctcgctgcct tcaccatgtt ctcccgcttc ctgaacaagc
1861   agccatactg atgaccacag caaccagctc cacggcctga tgcagcccct cccagcctct
1921   cccgctagga gagtcctctt ctaagcaaag tgcccctgca ggccgggttc tgccgccagg
1981   actgcccctt tcccagagcc ctgtacatcc cagactgggc ccagggtctc ccatagacag
2041   cctgggggca agttagcact ttattcccgc agcagttcct gaatggggtg gcctggcccc
2101   ttctctgctt aaagaatgcc ctttatgatg cactgattcc atcccaggaa cccaacagag
2161   ctcaggacag cccacaggga ggtggtggac ggactgtaat tgatagattg attatggaat
2221   taaattgggt acagcttcaa aaaaaaaaaa aaaa
```

SEQ ID NO: 2 Human PPCA Polypeptide, variant 1 protein

```
MTSSPRAPPGEQGRGGAEMIRAAPPPLFLLLLLLLLVSWASRG
EAAPDQDEIQRLPGLAKQPSFRQYSGYLKGSGSKHLHYWFVESQKDPENSPVVLWLNG
GPGCSSLDGLLTEHGPFLVQPDGVTLEYNPYSWNLIANVLYLESPAGVGFSYSDDKFY
AINDTEVAQSNFEALQDFFRLFPEYKNNKLFLIGESYAGIYIPTLAVLVMQDPSMNLQ
GLAVGNGLSSYEQNDNSLVYFAYYHGLLGNRLWSSLQTHCCSQNKCNFYDNKDLECVT
NLQEVARIVGNSGLNIYNLYAPCAGGVPSHFRYEKDIVVVQDLGNIFIRLPLKRMWHQ
```

-continued

SEQUENCE LISTING

```
ALLRSGDKVRMDPPCINTTAASTYLNNPYVRKALNIPEQLPQWDMCNFLVNLQYRRLY
RSMNSQYLKLLSSQKYQILLYNGDVDMACNFMGDEWFVDSLNQKMEVQRRPWLVKYGD
SGEQIAGFVKEFSHIAFLTIKGAGHMVPTDKPLAAFTMFSRFLNKQPY

SEQ ID NO: 3 Human NEU1 mRNA
   1   gagctacttg aagaccaatt agagtccggg aagcgcggcg gggcctccag accggggcgg
  61   gcttaagggt gacatctgcg ctttaaaggg tccgggtcag ctgactcccg actctgtgga
 121   gtctagctgc cagggtcgcg gcagctgcgg ggagagatga ctggggagcg acccagcacg
 181   gcgctcccgg acagacgctg ggggccgcgg attctgggct tctggggagg ctgtagggtt
 241   tgggtgtttg ccgcgatctt cctgctgctg tctctggcag cctcctggtc caaggctgag
 301   aacgacttcg gtctggtgca gccgctggtg accatggagc aactgctgtg ggtgagcggg
 361   agacagatcg gctcagtgga caccttccgc atcccgctca tcacagccac tccgcggggc
 421   actcttctcg cctttgctga ggcgaggaaa atgtcctcat ccgatgaggg ggccaagttc
 481   atcgccctgc ggaggtccat ggaccagggc agcacatggt ctcctacagc gttcattgtc
 541   aatgatgggg atgtccccga tgggctgaac cttggggcag tagtgagcga tgttgagaca
 601   ggagtagtat ttcttttcta ctccctttgt gctcacaagg ccggctgcca ggtggcctct
 661   accatgttgg tatggagcaa ggatgatggt gtttcctgga gcacaccccg gaatctctcc
 721   ctggatattg gcactgaagt gtttgcccct ggaccgggct ctggtattca gaaacagcgg
 781   gagccacgga agggccgcct catcgtgtgt ggccatggga cgctggagcg ggacggagtc
 841   ttctgtctcc tcagcgatga tcatggtgcc tcctggcgct acggaagtgg ggtcagcggc
 901   atcccctacg gtcagcccaa gcaggaaaat gatttcaatc ctgatgaatg ccagccctat
 961   gagctcccag atggctcagt cgtcatcaat gcccgaaacc agaacaacta ccactgccac
1021   tgccgaattg tcctccgcag ctatgatgcc tgtgatacac taaggccccg tgatgtgacc
1081   ttcgaccctg agctcgtgga ccctgtggta gctgcaggag ctgtagtcac cagctccggc
1141   attgtcttct tctccaaccc agcacatcca gagttccgag tgaacctgac cctgcgatgg
1201   agcttcagca atggtacctc atggcggaaa gagacagtcc agctatggcc aggccccagt
1261   ggctattcat ccctggcaac cctggagggc agcatggatg gagaggagca ggccccccag
1321   ctctacgtcc tgtatgagaa aggccggaac cactacacag agagcatctc cgtggccaaa
1381   atcagtgtct atgggacact ctgagctgtg ccactgccac aggggtattc tgccttcagg
1441   actctgcctt caggaacacg ggtctgtaga gggtctgctg gagacgcctg aaagacagtt
1501   ccatcttcct ttagactcca gccttggcaa aatcaccttc cctttaccag ggaaatcact
1561   tcctttagga ctgaaagcta ggcgtcctct cccacaaaaa agtcctgccc tcatctgaga
1621   atactgtctt tccatatggc taagtgtggc cccaccaccc tctctgccct cccgggacat
1681   tgattggtcc tgtcttgggc aggtctagtg agctgtagaa ttgaatcaat gtgaactcag
1741   ggaactgggg aaggctgagc ctcctctttg gtgttgcggt aagataaccg acagggctgg
1801   tgaaagtccc cagatggcag gatatttggt ttcagagtaa ggactaggtg caccaccatg
1861   actgactatc aatcaaaatg tttgtaactt aaaatttta atgaaggata atgaatattt
1921   gtagagtctc tatggttctg tcaatgcaca tcttcgtgtc tgttttcctc atgtatcctt
1981   gtgagcctgg gtgagttctg gggagagacc tgatgtgcgt actgcctgtg aaaatctgac
2041   tttggcaaat caaatcctct tttccttttg aaaaaaaaaa aaaaaaaa

SEQ ID NO: 4 Human NEU1 Polypeptide
        10         20         30         40         50
MTGERPSTAL PDRRWGPRIL GFWGGCRVWV FAAIFLLLSL AASWSKAEND
        60         70         80         90        100
FGLVQPLVTM EQLLWVSGRQ IGSVDTFRIP LITATPRGTL LAFAEARKMS
       110        120        130        140        150
SSDEGAKFIA LRRSMDQGST WSPTAFIVND GDVPDGLNLG AVVSDVETGV
       160        170        180        190        200
VFLFYSLCAH KAGCQVASTM LVWSKDDGVS WSTPRNLSLD IGTEVFAPGP
       210        220        230        240        250
GSGIQKQREP RKGRLIVCGH GTLERDGVFC LLSDDHGASW RYGSGVSGIP
       260        270        280        290        300
YGQPKQENDF NPDECQPYEL PDGSVVINAR NQNNYHCHCR IVLRSYDACD
       310        320        330        340        350
TLRPRDVTFD PELVDPVVAA GAVVTSSGIV FFSNPAHPEF RVNLTLRWSF
       360        370        380        390        400
SNGTSWRKET VQLWPGPSGY SSLATLEGSM DGEEQAPQLY VLYEKGRNHY
       410
TESISVAKIS VYGTL

SEQ ID NO: 5 Human TPP1 mRNA
   1   ggtggtggaa tatagagctc atgtgatccg tcacatgaca gcagatccgc ggaagggcag
  61   aatgggactc caagcctgcc tcctagggct ctttgccctc atcctctctg gcaaatgcag
 121   ttacagcccg gagcccgacc agcggaggac gctgccccca ggctgggtgt ccctgggccg
 181   tgcggaccct gaggaagagc tgagtctcac ctttgccctg agacagcaga atgtggaaag
 241   actctcggag ctggtgcagg ctgtgtcgga tcccagctct cctcaatacg gaaaatacct
 301   gaccctagag aatgtggctg atctggtgag gccatcccca ctgaccctcc acacggtgca
 361   aaaatggctc ttggcagccg gagcccagaa gtgccattct gtgatcacac aggactttct
 421   gacttgctgg ctgagcatcc gacaagcaga gctgctgctc cctggggctg agtttcatca
 481   ctatgtggga ggacctacgg aaacccatgt tgtaaggtcc ccacatccct accagcttcc
 541   acaggccttg gccccccatg tggactttgt gggggggactg caccgttttc ccccaacatc
 601   atccctgagg caacgtcctg agccgcaggt gacagggact gtaggcctgc atctgggggt
 661   aaccccctct gtgatccgta agcgatacaa cttgacctca caagacgtgg gctctggcac
```

-continued

SEQUENCE LISTING

```
 721   cagcaataac agccaagcct gtgcccagtt cctggagcag tatttccatg actcagacct
 781   ggctcagttc atgcgcctct tcggtggcaa ctttgcacat caggcatcag tagcccgtgt
 841   ggttggacaa cagggccggg gccgggccgg gattgaggcc agtctagatg tgcagtacct
 901   gatgagtgct ggtgccaaca tctccacctg ggtctacagt agccctggcc ggcatgaggg
 961   acaggagccc ttcctgcagt ggctcatgct gctcagtaat gagtcagccc tgccacatgt
1021   gcatactgtg agctatggag atgatgagga ctccctcagc agcgcctaca tccagcgggt
1081   caacactgag ctcatgaagg ctgccgctcg gggtctcacc ctgctctttcg cctcaggtga
1141   cagtggggcc gggtgttggt ctgtctctgg aagacaccag ttccgcccta ccttccctgc
1201   ctccagcccc tatgtcacca cagtgggagg cacatccttc caggaacctt tcctcatcac
1261   aaatgaaatt gttgactata tcagtggtgg tggcttcagc aatgtgttcc cacggccttc
1321   ataccaggag gaagctgtaa cgaagttcct gagctctagc ccccacctgc caccatccag
1381   ttacttcaat gccagtggcc gtgcctaccc agatgtggct gcactttctg atggctactg
1441   ggtggtcagc aacagagtgc ccattccatg ggtgtccgga acctcggcct ctactccagt
1501   gtttgggggg atcctatcct tgatcaatga gcacaggatc cttagtggcc gccccctct
1561   tggctttctc aacccaaggc tctaccagca gcatggggca ggactctttg atgtaacccg
1621   tggctgccat gagtcctgtc tggatgaaga ggtagagggc cagggtttct gctctggtcc
1681   tggctgggat cctgtaacag gctggggaac acccaacttc ccagctttgc tgaagactct
1741   actcaacccc tgacccttcc ctatcaggag agatggcttg tcccctgccc tgaagctggc
1801   agttcagtcc cttattctgc cctgttggaa gccctgctga accctcaact attgactgct
1861   gcagacagct tatctcccta accctgaaat gctgtgagct tgacttgact cccaacccta
1921   ccatgctcca tcatactcag gtctccctac tcctgcctta gattcctcaa taagatgctg
1981   taactagcat tttttgaatg cctctccctc cgcatctcat ctttctcttt tcaatcaggc
2041   ttttccaaag ggttgtatac agactctgtg cactatttca cttgatattc attccccaat
2101   tcactgcaag gagacctcta ctgtcaccgt ttactctttc ctaccctgac atccagaaac
2161   aatggcctcc agtgcatact tctcaatctt tgctttatgg cctttccatc atagttgccc
2221   actccctctc cttacttagc ttccaggtct taacttctct gactactctt gtcttcctct
2281   ctcatcaatt tctgcttctt catggaatgc tgaccttcat tgctccattt gtagattttt
2341   gctcttctca gtttactcat tgtcccctgg aacaaatcac tgacatctac aaccattacc
2401   atctcactaa ataagacttt ctatccaata atgattgata cctcaaatgt aagatgcgtg
2461   atactcaaca tttcatcgtc caccttccca accccaaaca attccatctc gtttcttctt
2521   ggtaaatgat gctatgcttt ttccaaccaa gccagaaacc tgtgtcatct tttcacccca
2581   ccttcaatca acaagtcctc aatcaacaag tcctactgac tgcacatctt aaatatatct
2641   ttatcagtcc acaagtcctt ccaattatat ttcccaagta tatctagaac ttatccactt
2701   atatccccac tgctactacc ttagtttagg gctatattct cttgaaaaaa agtgtcctta
2761   cttcctgcca atccccaagt catcttccag agtaaaatgc aaatcccatc aggccacttg
2821   gatgaaaacc cttcaaggat tactggatag aattcaggct ttcccctcca gcccccaatc
2881   atagctcaca aaccttcctt gctatttgtt cttaagtaaa aaatcatttt tcctcctccc
2941   tccccaaacc ccaaggaact ctcactcttg ctcaagctgt tccgtcccct taccacccct
3001   gatacaactg ccaggttaat ttccagaatt cttgcaagac tcagttcaga agtcaccttc
3061   tttcgtgaat gttttgattc cctgaggcta ctttatttttg gtatggctga aaaatcctag
3121   attttctaaa caaaacctgt ttgaatcttg gttctgatat ggactaggag agagactggg
3181   tcaagtaagc ttatctccct gaggctgttt cctcgtctgt taagtgtgaa tatcaatacc
3241   tgcctttcat aatcaccagg gaataaagtg gaataatgtt gataacagtg cttggcacct
3301   ggaagtaggt ggcagatgtt aacgcccttc ctcccttgca ctgcgccccc tgtgcctacc
3361   tctagcattg taacgaccac gtagtattga aatggccagt ttacttgtct gccttccttt
3421   ccaagaccgt tggtgcctag aggactagaa tcgtgtccta tttaactttg tgttcccagg
3481   tcctagctca ggagttggca aataagaatt aaatgtctgc tacaccgaaa accaaaaaaa
```

SEQ ID NO: 6 Human TPP1 Polypeptide

```
        10         20         30         40         50
MGLQACLLGL FALILSGKCS YSPEPDQRRT LPPGWVSLGR ADPEEELSLT
        60         70         80         90        100
FALRQQNVER LSELVQAVSD PSSPQYGKYL TLENVADLVR PSPLTLHTVQ
       110        120        130        140        150
KWLLAAGAQK CHSVITQDFL TCWLSIRQAE LLLPGAEFHH YVGGPTETHV
       160        170        180        190        200
VRSPHPYQLP QALAPHVDFV GGLHRFPPTS SLRQRPEPQV TGTVGLHLGV
       210        220        230        240        250
TPSVIRKRYN LTSQDVGSGT SNNSQACAQF LEQYFHDSDL AQFMRLFGGN
       260        270        280        290        300
FAHQASVARV VGQQGRGRAG IEASLDVQYL MSAGANISTW VYSSPGRHEG
       310        320        330        340        350
QEPFLQWLML LSNESALPHV HTVSYGDDED SLSSAYIQRV NTELMKAAAR
       360        370        380        390        400
GLTLLFASGD SGAGCWSVSG RHQFRPTFPA SSPYVTTVGG TSFQEPFLIT
       410        420        430        440        450
NEIVDYISGG GFSNVFPRPS YQEEAVTKFL SSSPHLPPSS YFNASGRAYP
       460        470        480        490        500
DVAALSDGYW VVSNRVPIPW VSGTSASTPV FGGILSLINE HRILSGRPPL
       510        520        530        540        550
GFLNPRLYQQ HGAGLFDVTR GCHESCLDEE VEGQGFCSGP GWDPVTGWGT
       560
PNFPALLKTL LNP
```

-continued

SEQUENCE LISTING

SEQ ID NO: 7 Human Cathepsin B mRNA, variant 1

```
   1   ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg
  61   ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct
 121   ctcggctgca gcgctgggtg gatctaggat ccggcttcca acatgtggca gctctgggcc
 181   tccctctgct gcctgctggt gttggccaat gcccggagca ggccctcttt ccatcccctg
 241   tcggatgagc tggtcaacta tgtcaacaaa cggaatacca cgtggcaggc cgggcacaac
 301   ttctacaacg tggacatgag ctacttgaag aggctatgtg gtaccttcct gggtgggccc
 361   aagccacccc agagagttat gtttaccgag gacctgaagc tgcctgcaag cttcgatgca
 421   cgggaacaat ggccacagtg tcccaccatc aaagagatca gagaccaggg ctcctgtggc
 481   tcctgctggg ccttcggggc tgtggaagcc atctctgacc ggatctgcat ccacaccaat
 541   gcgcacgtca gcgtggaggt gtcggcggag gacctgctca catgctgtgg cagcatgtgt
 601   ggggacggct gtaatggtgg ctatcctgct gaagcttgga acttctggac aagaaaaggc
 661   ctggtttctg gtggcctcta tgaatcccat gtagggtgca gaccgtactc catccctccc
 721   tgtgagcacc acgtcaacgg ctcccggccc ccatgcacgg gggagggaga taccccccaag
 781   tgtagcaaga tctgtgagcc tggctacagc ccgacctaca aacaggacaa gcactacgga
 841   tacaattcct acagcgtctc caatagcgag aaggacatca tggccgagat ctacaaaaac
 901   ggccccgtgg agggagcttt ctctgtgtat tcggacttcc tgctctacaa gtcaggagtg
 961   taccaacacg tcaccggaga gatgatgggt ggccatgcca tccgcatcct gggctgggga
1021   gtggagaatg gcacacccta ctggctggtt gccaactcct ggaacactga ctggggtgac
1081   aatggcttct ttaaaatact cagaggacag gatcactgtg gaatcgaatc agaagtggtg
1141   gctggaattc cacgcaccga tcagtactgg gaaaagatct aatctgccgt gggcctgtcg
1201   tgccagtcct gggggcgaga tcggggtaga aatgcatttt attctttaag ttcacgtaag
1261   atacaagttt cagacagggt ctgaaggact ggattggcca aacatcagac ctgtcttcca
1321   aggagaccaa gtcctggcta catcccagcc tgtggttaca gtgcagacag gccatgtgag
1381   ccaccgctgc cagcacagag cgtccttccc cctgtagact agtgccgtag ggagtacctg
1441   ctgccccagc tgactgtggc cccctccgtg atccatccat ctccagggag caagacagag
1501   acgcaggaat ggaaagcgga gttcctaaca ggatgaaagt tcccccatca gttcccccag
1561   tacctccaag caagtagctt tccacatttg tcacagaaat cagaggagag acggtgttgg
1621   gagccctttg gagaacgcca gtctcccagg ccccctgcat ctatcgagtt tgcaatgtca
1681   caacctctct gatcttgtgc tcagcatgat tctttaatag aagttttatt ttttcgtgca
1741   ctctgctaat catgtgggtg agccagtgga acagcgggag acctgtgcta gttttacaga
1801   ttgcctcctt atgacgcggc tcaaaaggaa accaagtggt caggagttgt ttctgaccca
1861   ctgatctcta ctaccacaag gaaaatagtt taggagaaac cagcttttac tgtttttgaa
1921   aaattacagc ttcaccctgt caagttaaca aggaatgcct gtgccaataa aagttttctc
1981   caacttgaag tctactctga tgggatctca gatcctttgt cactgcctat agacttgtag
2041   ctgctgtctc tctttgtccc tgcagagaat cacgtcctgg aactgcatgt tcttgcgact
2101   cttgggactt catcttaact tctcgctgcc ccagccatgt tttcaaccat ggcatccctc
2161   ccccaattag ttccctgtca tcctcgtcaa ccttctctgt aagtgcctgg taagcttgcc
2221   cttgcttaag aactcaaaac atagctgtgc tctatttttt tgttgttgtt gtgactgaca
2281   gagtgagatt ccgtctccca ggctggagtg cagtggcgcc ttctcagctc actgcaacct
2341   gcagcctcct agattcaagc gattctcctg cttcagcctt ccgagtagct gggatgacag
2401   gcactcacca atatgcctgg gtaatttttg tatttttaag tacatacagg atttcaccat
2461   gttggccagg ctagtttcaa actcccggcc tcaggtggtc tgcctgcctc agcctcccaa
2521   agtgttggga ttacaggcgt gagccactgg gccctgcctg tatttttttat cagccacaaa
2581   tccagcaaca agctgaggat tcagctcata aaacaggctt ggtgtcttgg tgatctcaca
2641   taaccaagat gctaccccgt ggggaaccac atccccctgg atgccctcca gccttggttt
2701   gggctggagt cagggcctgt atacagtatt ttgaatttgt atgccactgg tttgcattgc
2761   tggtcaggaa ctctagtgct ttgcatagcc ctggtttaga aacatgttat agcagttctt
2821   ggtatagagc aaactagaag aaccagcaat cattccactg tcctgccaag gtacacctca
2881   gtactcccct tcccaactga agtggtatga ggctagctct ttccaaaagc attcaagttt
2941   ggcttctgat gtgactcaga atttaggaac cagatgctag atcaaataag ctctgaaaat
3001   ctgaggaaca ttgtaggaaa ggtttgttaa gcatctctta agtgccatga tgagcataac
3061   agccggccgt cgtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggagga
3121   tgacaaggtc aggagttcaa gaccagcctg gccaacatgc tgaaacctca cctctactaa
3181   aaatacaaaa attagctggg catggtggca catgcctgta atcccagcta cttgggaggc
3241   tgaggcagga gaatcgcttg aacccgggag gcggaggttg cagtgagcca agacagtgcc
3301   agtgcactcc agcctcggtg acagcgcaag gctccgtctc aataattaaa aaaaaaaaaa
3361   aaaaaaaaaa ggccgggcgc agtggctcaa gcctgtaatc ccagcacttt gggaggctga
3421   ggcgggcaga tcacctgagg tcaggagttt tgagatcagc cttggcaaca cggtgaaacc
3481   ccatctctac taaaaataca aaattagcca agcatgctgg cacatgcctg taatcccagc
3541   tactcgggag gctgaggtac gagaatcgct tgaacctggg aggcagagga tgcagtgagc
3601   cgagatcacg ccattgcact ccagcctggg ggacaagagt gaatctgtgt ctcaccaaaa
3661   aaaaaaagaa aaagaaagat gcttaacaaa ggttaccata agccacaaat tcataaccac
3721   ttatccttcc agtttcaagt agaatatatt cataacctca ataaagttct ccctgctccc
3781   aaa
```

SEQ ID NO: 8 Human Cathepsin B Polypeptide, variant 1

```
MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT
YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI
```

-continued

SEQUENCE LISTING

SEQ ID NO: 9 Human Cathepsin K mRNA

```
   1   acacatgctg catacacaca gaaacactgc aaatccactg cctccttccc tcctccctac
  61   ccttccttct ctcagcattt ctatccccgc ctcctcctct tacccaaatt ttccagccga
 121   tcactggagc tgacttccgc aatcccgatg gaataaatct agcacccctg atggtgtgcc
 181   cacactttgc tgccgaaacg aagccagaca acagatttcc atcagcagga tgtgggggct
 241   caaggttctg ctgctacctg tggtgagctt tgctctgtac cctgaggaga tactggacac
 301   ccactgggag ctatggaaga agacccacag gaagcaatat aacaacaagg tggatgaaat
 361   ctctcggcgt ttaatttggg aaaaaaacct gaagtatatt tccatccata accttgaggc
 421   ttctcttggt gtccatacat atgaactggc tatgaaccac ctgggggaca tgaccagtga
 481   agaggtggtt cagaagatga ctggactcaa agtacccctg tctcattccc gcagtaatga
 541   cacccttat atcccagaat gggaaggtag agccccagac tctgtcgact atcgaaagaa
 601   aggatatgtt actcctgtca aaaatcaggg tcagtgtggt tcctgttggg cttttagctc
 661   tgtgggtgcc ctggagggcc aactcaagaa gaaaactggc aaactcttaa atctgagtcc
 721   ccagaaccta gtggattgtg tgtctgagaa tgatggctgt ggaggggggct acatgaccaa
 781   tgccttccaa tatgtgcaga agaaccgggg tattgactct gaagatgcct acccatatgt
 841   gggacaggaa gagagttgta tgtacaaccc aacaggcaag gcagctaaat gcagagggta
 901   cagagagatc cccgagggga atgagaaagc cctgaagagg gcagtggccc gagtgggacc
 961   tgtctctgtg gccattgatg caagcctgac ctccttccag ttttacagca aaggtgtgta
1021   ttatgatgaa agctgcaata gcgataatct gaaccatgcg gttttggcag tgggatatgg
1081   aatccagaag ggaaacaagc actggataat taaaaacagc tggggagaaa actggggaaa
1141   caaaggatat atcctcatgg ctcgaaataa gaacaacgcc tgtggcattg ccaacctggc
1201   cagcttcccc aagatgtgac tccagccagc caaatccatc ctgctcttcc atttcttcca
1261   cgatggtgca gtgtaacgat gcactttgga agggagttgg tgtgctattt ttgaagcaga
1321   tgtggtgata ctgagattgt ctgttcagtt tccccatttg tttgtgcttc aaatgatcct
1381   tcctactttg cttctctcca cccatgacct ttttcactgt ggccatcagg actttccctg
1441   acagctgtgt actcttaggc taagagatgt gactacagcc tgcccctgac tgtgttgtcc
1501   cagggctgat gctgtacagg tacaggctgg agattttcac ataggttaga ttctcattca
1561   cgggactagt tagctttaag caccctagag gactagggta atctgacttc tcacttccta
1621   agttcccttc tatatcctca aggtagaaat gtctatgttt tctactccaa ttcataaatc
1681   tattcataag tctttggtac aagtttacat gataaaaaga aatgtgattt gtcttccctt
1741   ctttgcactt ttgaaataaa gtatttatct cctgtctaca gtttaataaa tagcatctag
1801   tacacattca aaaaaaaaaa aaaaa
```

SEQ ID NO: 10 Human Cathepsin K Polypeptide

```
        10          20          30          40          50
MWGLKVLLLP VVSFALYPEE ILDTHWELWK KTHRKQYNNK VDEISRRLIW
        60          70          80          90         100
EKNLKYISIH NLEASLGVHT YELAMNHLGD MTSEEVVQKM TGLKVPLSHS
       110         120         130         140         150
RSNDTLYIPE WEGRAPDSVD YRKKGYVTPV KNQGQCGSCW AFSSVGALEG
       160         170         180         190         200
QLKKKTGKLL NLSPQNLVDC VSENDGCGGG YMTNAFQYVQ KNRGIDSEDA
       210         220         230         240         250
YPYVGQEESC MYNPTGKAAK CRGYREIPEG NEKALKRAVA RVGPVSVAID
       260         270         280         290         300
ASLTSFQFYS KGVYYDESCN SDNLNHAVLA VGYGIQKGNK HWIIKNSWGE
       310         320
NWGNKGYILM ARNKNNACGI ANLASFPKM
```

SEQ ID NO: 11 Human Cathepsin L mRNA, variant 1

```
   1   ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc
  61   agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac
 121   cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag
 181   gacgcggtcg agtaggtgtg caccagccct ggcaacgaga gcgtctaccc cgaactctgc
 241   tggccttgag gtggggaagc cggggagggc agttgaggac cccgcggagg cgcgtgactg
 301   gttgagcggg caggccagcc tccgagccgg gtggacacag gttttaaaac atgaatccta
 361   cactcatcct tgctgccttt tgcctgggaa ttgcctcagc tactctaaca tttgatcaca
 421   gtttagaggc acagtggacc aagtggaagg cgatgcacaa cagattatac ggcatgaatg
 481   aagaaggatg gaggagagca gtgtgggaga agaacatgaa gatgattgaa ctgcacaatc
 541   aggaatacag ggaagggaaa cacagcttca caatggccat gaacgccttt ggagacatga
 601   ccagtgaaga attcaggcag gtgatgaatg gctttcaaaa ccgtaagccc aggaagggga
 661   aagtgttcca ggaacctctg ttttatgagg cccccagatc tgtggattgg agagagaaag
 721   gctacgtgac tcctgtgaag aatcagggtc agtgtggttc ttgttgggct tttagtgcta
 781   ctggtgctct tgaaggacag atgttccgga aaactgggag gcttatctca ctgagtgagc
 841   agaatctggt agactgctct gggcctcaag gcaatgaagg ctgcaatggt ggcctaatgg
 901   attatgcttt ccagtatgtt caggataatg gaggcctgga ctctgaggaa tcctatccat
 961   atgaggcaac agaagaatcc tgtaagtaca atcccaagta ttctgttgct aatgacaccg
1021   gctttgtgga catccctaag caggagaagg ccctgatgaa ggcagttgca actgtggggc
1081   ccatttctgt tgctattgat gcaggtcatg agtccttcct gttctataaa gaaggcattt
1141   attttgagcc agactgtagc agtgaagaca tggatcatgg tgtgctggtg gttggctacg
1201   gatttgaaag cacagaatca gataacaata aatattggct ggtgaagaac agctggggtg
1261   aagaatgggg catgggtggc tacgtaaaga tggccaaaga ccggagaaac cattgtggaa
1321   ttgcctcagc agccagctac cccactgtgt gagctggtgg acggtgatga ggaaggactt
```

-continued

SEQUENCE LISTING

1381 gactggggat ggcgcatgca tgggaggaat tcatcttcag tctaccagcc cccgctgtgt
1441 cggatacaca ctcgaatcat tgaagatccg agtgtgattt gaattctgtg atattttcac
1501 actggtaaat gttacctcta ttttaattac tgctataaat aggtttatat tattgattca
1561 cttactgact ttgcattttc gtttttaaaa ggatgtataa atttttacct gtttaaataa
1621 aatttaattt caaatgtagt ggtggggctt ctttctattt ttgatgcact gaatttttgt
1681 gtaataaaga acataattgg gctctaagcc ataaaaaaaa aaaaaaaaaa

SEQ ID NO: 12 Human Cathepsin L Polypeptide, variant 1
MNPTLILAAFCLGIASATLIFDHSLEAQWTKWKAMHNRLYGMNE
EGWRRAVWEKNMKMIELHNQEYREGKHSFTMAMNAFGDMISEEFRQVMNGFQNRKPRK
GKVFQEPLFYEAPRSVDWREKGYVTPVKNQGQCGSCWAFSATGALEGQMFRKTGRLIS
LSEQNLVDCSGPQGNEGCNGGLMDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYS
VANDTGFVDIPKQEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDH
GVLVVGYGFESTESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV

SEQ ID NO: 13
DXXLL

SEQ ID NO: 14
[DE]XXXL[LI]

SEQ ID NO: 15
YXXØ

SEQ ID NO: 16, MPR300/CI-MPR
SFHDDSDEDLL

SEQ ID NO: 17, MPR46/CD-MPR
EESEERDDHLL

SEQ ID NO: 18 Sortilin
GYHDDSDEDLL

SEQ ID NO: 19 SorLA/SORL1
ITGFSDDVPMV

SEQ ID NO: 20 GGA1 (1)
ASVSLLDDELM

SEQ ID NO: 21 GGA1 (2)
ASSGLDDLDLL

SEQ ID NO: 22, GGA2
VQNPSADRNLL

SEQ ID NO: 23, GGA3
NALSWLDEELL

SEQ ID NO: 24, LIMP-II
DERAPLI

SEQ ID NO: 25, NPC1
TERERLL

SEQ ID NO: 26, Mucolipin-1
SETERLL

SEQ ID NO: 27, Sialin
TDRTPLL

SEQ ID NO: 28, GLUT8
EETQPLL

SEQ ID NO: 29, Invariant chain (Ii) (1)
DDQRDLI

SEQ ID NO: 30, Invariant chain (Ii) (2)
NEQLPML

SEQ ID NO: 31, LAMP-1
GYQTI

SEQ ID NO: 32, LAMP-2A
GYEQF

-continued

SEQUENCE LISTING

SEQ ID NO: 33, LAMP-2B
GYQTL

SEQ ID NO: 34, LAMP-2C
GYQSV

SEQ ID NO: 35, CD63
GYEVM

SEQ ID NO: 36, CD68
AYQAL

SEQ ID NO: 37, Endolyn
NYHTL

SEQ ID NO: 38, DC-LAMP
GYQRI

SEQ ID NO: 39, Cystinosin
GYDQL

SEQ ID NO: 40, Sugar phosphate exchanger 2
GYKEI

SEQ ID NO: 41, acid phosphatase
GYRHV

SEQ ID NO: 42, Human PPCA, variant 2 mRNA

```
   1   agagtgcacc cgaatccacg ggctcggagg cagcagccat ctctcggcca tagggcaggc
  61   cagctggcgc cgggggctat tttgggcggc gggcaatgat ggtgaccgca aggcgacctt
 121   gtaaggcatt tccccctga ctcccttccc cgagcctctg cccgggggtc ctagcgccgc
 181   tttctcagcc atcccgccta caacttagcc gtccacaaca ggatcatctg atcgcgtgcg
 241   cccgggctac gatctgcgag gcccgcggac cttgacccgg cattgaccgc caccgccccc
 301   caggtccgta gggaccaaag aaggggcggg aggaagactg tcacgtggcg ccggagttca
 361   cgtgactcgt acacatgact tccagtcccc gggcgcctcc tggagagcaa ggacgcgggg
 421   gagcagaggt gagctggcac cggaggctgg aggggatccc cgagcccggg atcgatgatc
 481   cgagccgcgc cgccgccgct gttcctgctg ctgctgctgc tgctgctgct agtgtcctgg
 541   gcgtcccgag gcgaggcagc ccccgaccag gacgagatcc agcgcctccc cgggctggcc
 601   aagcagccgt ctttccgcca gtactccggc tacctcaaag gctccggctc caagcacctc
 661   cactactggt ttgtggagtc ccagaaggat cccgagaaca gccctgtggt gctttggctc
 721   aatgggggtc ccggctgcag ctcactagat gggctcctca cagagcatgg ccccttcctg
 781   gtccagccag atggtgtcac cctggagtac aacccctatt cttggaatct gattgccaat
 841   gtgttatacc tggagtcccc agctggggtg ggcttctcct actccgatga caagttttat
 901   gcaactaatg acactgaggt cgcccagagc aattttgagg cccttcaaga tttcttccgc
 961   ctctttccgg agtacaagaa caacaaactt ttcctgaccg gggagagcta tgctggcatc
1021   tacatcccca ccctggccgt gctggtcatg caggatccca gcatgaacct tcaggggctg
1081   gctgtgggca atggactctc ctcctatgag cagaatgaca actccctggt ctactttgcc
1141   tactaccatg gccttctggg gaacaggctt tggtcttctc tccagaccca ctgctgctct
1201   caaaacaagt gtaacttcta tgacaacaaa gacctggaat gcgtgaccaa tcttcaggaa
1261   gtggcccgca tcgtgggcaa ctctggcctc aacatctaca atctctatgc cccgtgtgct
1321   ggaggggtgc ccagccattt taggtatgag aaggacactg ttgtggtcca ggatttgggc
1381   aacatcttca ctcgcctgcc actcaagcgg atgtggcatc aggcactgct gcgctcaggg
1441   gataaagtgc gcatggaccc cccctgcacc aacacaacag ctgcttccac ctacctcaac
1501   aacccgtacg tgcggaaggc cctcaacatc ccggagcagc tgccacaatg ggacatgtgc
1561   aactttctgg taaacttaca gtaccgccgt ctctaccgaa gcatgaactc ccagtatctg
1621   aagctgctta gctcacagaa ataccagatc ctattatata atggagatgt agacatggcc
1681   tgcaatttca tgggggatga gtggtttgtg gattccctca accagaagat ggaggtgcag
1741   cgccggccct ggttagtgaa gtacggggac agcggggagc agattgccgg cttcgtgaag
1801   gagttctccc acatcgcctt tctcacgatc aagggcgccg gccacatggt tcccaccgac
1861   aagcccctcg ctgccttcac catgttctcc cgcttcctga acaagcagcc atactgatga
1921   ccacagcaac cagctccacg gcctgatgca gcccctccca gcctctcccg ctaggagagt
1981   cctcttctaa gcaaagtgcc cctgcaggcc gggttctgcc gccaggactg ccccttccc
2041   agagccctgt acatcccaga ctgggcccag ggtctcccat agacagcctg ggggcaagtt
2101   agcactttat tcccgcagca gttcctgaat ggggtggcct ggccccttct ctgcttaaag
2161   aatgcccttt atgatgcact gattccatcc caggaaccca acagagctca ggacagccca
2221   cagggaggtg gtggacggac tgtaattgat agattgatta tggaattaaa ttgggtacag
2281   cttcaaaaaa aaaaaaaaaa
```

SEQ ID NO: 43, Human PPCA, variant 2 protein

```
        10         20         30         40         50
MIRAAPPPLF LLLLLLLLLV SWASRGEAAP DQDEIQRLPG LAKQPSFRQY
        60         70         80         90        100
```

-continued

SEQUENCE LISTING

```
SGYLKGSGSK HLHYWFVESQ KDPENSPVVL WLNGGPGCSS LDGLLTEHGP
       110         120         130         140         150
FLVQPDGVTL EYNPYSWNLI ANVLYLESPA GVGFSYSDDK FYATNDTEVA
       160         170         180         190         200
QSNFEALQDF FRLFPEYKNN KLFLTGESYA GIYIPTLAVL VMQDPSMNLQ
       210         220         230         240         250
GLAVGNGLSS YEQNDNSLVY FAYYHGLLGN RLWSSLQTHC CSQNKCNFYD
       260         270         280         290         300
NKDLECVTNL QEVARIVGNS GLNIYNLYAP CAGGVPSHFR YEKDTVVVQD
       310         320         330         340         350
LGNIFTRLPL KRMWHQALLR SGDKVRMDPP CTNTTAASTY LNNPYVRKAL
       360         370         380         390         400
NIPEQLPQWD MCNFLVNLQY RRLYRSMNSQ YLKLLSSQKY QILLYNGDVD
       410         420     430 440 450
MACNFMGDEW FVDSLNQKME VQRRPWLVKY GDSGEQIAGF VKEFSHIAFL
       460         470 480
TIKGAGHMVP TDKPLAAFTM FSRFLNKQPY
```

SEQ ID NO: 44, Human PPCA, variant 3 mRNA

```
   1   agagtgcacc cgaatccacg ggctcggagg cagcagccat ctctcggcca tagggcaggc
  61   cagctggcgc cgggggctat tttgggcggc gggcaatgat ggtgaccgca aggcgacctt
 121   gtaaggcatt tccccccctga ctcccttccc cgagcctctg cccgggggtc ctagcgccgc
 181   tttctcagcc atcccgccta caacttagcc gtccacaaca ggatcatctg atcgcgtgcg
 241   cccgggctac gatctgcgag gcccgcggac cttgacccgg cattgaccgc caccgccccc
 301   caggtccgta gggaccaaag aaggggcggg aggaagactg tcacgtggcg ccggagttca
 361   cgtgactcgt acacatgact tccagtcccc gggcgcctcc tggagagcaa ggacgcgggg
 421   gagcagagat gatccgagcc gcgccgccgc cgctgttcct gctgctgctg ctgctgctgc
 481   tgctagtgtc ctgggcgtcc cgaggcgagg cagcccccga ccaggacgag atccagcgcc
 541   tccccgggct ggccaagcag ccgtctttcc gccagtactc cggctacctc aaaggctccg
 601   gctccaagca cctccactac tggtttgtgg agtcccagaa ggatcccgag aacagccctg
 661   tggtgctttg gctcaatggg ggtcccggct gcagctcact agatgggctc ctcacagagc
 721   atggcccctt cctgattgcc aatgtgttat acctggagtc cccagctggg gtgggcttct
 781   cctactccga tgacaagttt tatgcaacta atgacactga ggtcgcccag agcaattttg
 841   aggcccttca agatttcttc cgcctctttc cggagtacaa gaacaacaaa cttttcctga
 901   ccggggagag ctatgctggc atctacatcc ccaccctggc cgtgctggtc atgcaggatc
 961   ccagcatgaa ccttcagggg ctggctgtgg gcaatggact ctcctcctat gagcagaatg
1021   acaactccct ggtctacttt gcctactacc atggccttct ggggaacagg ctttggtctt
1081   ctctccagac ccactgctgc tctcaaaaca agtgtaactt ctatgacaac aaagacctgg
1141   aatgcgtgac caatcttcag gaagtggccc gcatcgtggg caactctggc ctcaacatct
1201   acaatctcta tgccccgtgt gctggagggg tgcccagcca ttttaggtat gagaaggaca
1261   ctgttgtggt ccaggatttg ggcaacatct tcactcgcct gccactcaag cggatgtggc
1321   atcaggcact gctgcgctca ggggataaag tgcgcatgga ccccccctgc accaacacaa
1381   cagctgcttc cacctacctc aacaacccgt acgtgcggaa ggccctcaac atcccggagc
1441   agctgccaca atgggacatg tgcaactttc tggtaaactt acagtaccgc cgtctctacc
1501   gaagcatgaa ctcccagtat ctgaagctgc ttagctcaca gaaataccag atcctattat
1561   ataatggaga tgtagacatg gcctgcaatt tcatggggga tgagtggttt gtggattccc
1621   tcaaccagaa gatggaggtg cagcgccggc cctggttagt gaagtacggg gacagcgggg
1681   agcagattgc cggcttcgtg aaggagttct cccacatcgc ctttctcacg atcaagggcg
1741   ccggccacat ggttcccacc gacaagcccc tcgctgcctt caccatgttc tcccgcttcc
1801   tgaacaagca gccatactga tgaccacagc aaccagctcc acggcctgat gcagccctc
1861   ccagcctctc ccgctaggag agtcctcttc taagcaaagt gcccctgcag gccgggttct
1921   gccgccagga ctgccccctt cccagagccc tgtacatccc agactgggcc cagggtctcc
1981   catagacagc ctgggggcaa gttagcactt tattcccgca gcagttcctg aatggggtgg
2041   cctggcccct tctctgctta aagaatgccc tttatgatgc actgattcca tcccaggaac
2101   ccaacagagc tcaggacagc ccacagggag gtggtggacg gactgtaatt gatagattga
2161   ttatggaatt aaattgggta cagcttcaaa aaaaaaaaaa aaaaaaaa
```

SEQ ID NO: 45, Human PPCA, variant 3 protein

```
MTSSPRAPPGEQGRGGAEMIRAAPPPLFLLLLLLLLVSWASRG
EAAPDQDEIQRLPGLAKQPSFRQYSGYLKGSGSKHLHYWFVESQKDPENSPVVLWLNG
GPGCSSLDGLLTEHGPFLIANVLYLESPAGVGFSYSDDKFYATNDTEVAQSNFEALQD
FFRLFPEYKNNKLFLTGESYAGIYIPTLAVLVMQDPSMNLQGLAVGNGLSSYEQNDNS
LVYFAYYHGLLGNRLWSSLQTHCCSQNKCNFYDNKDLECVTNLQEVARIVGNSGLNIY
NLYAPCAGGVPSHFRYEKDTVVVQDLGNIFTRLPLKRMWHQALLRSGDKVRMDPPCTN
TTAASTYLNNPYVRKALNIPEQLPQWDMCNFLVNLQYRRLYRSMNSQYLKLLSSQKYQ
ILLYNGDVDMACNFMGDEWFVDSLNQKMEVQRRPWLVKYGDSGEQIAGFVKEFSHIAF
LTIKGAGHMVPTDKPLAAFTMFSRFLNKQPY
```

SEQ ID NO: 46 Human Cathepsin B mRNA, variant 2

```
   1   ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg
  61   ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct
 121   ctcggctgca gcgctgggct ggtgtgcagt ggtgcgacca cggctcacgg cagcctcagc
 181   cacccagatg taagcgatct ggttcccacc tcagcctccc gagtagtgtc ttcaggccta
 241   tggagagcag cttgcgtggg ctgggcctgc agtacctggt ttgcatagat gattggcagg
```

-continued

SEQUENCE LISTING

```
 301   tggatctagg atccggcttc caacatgtgg cagctctggg cctccctctg ctgcctgctg
 361   gtgttggcca atgcccggag caggccctct ttccatcccc tgtcggatga gctggtcaac
 421   tatgtcaaca aacggaatac cacgtggcag gccgggcaca acttctacaa cgtggacatg
 481   agctacttga agaggctatg tggtaccttc ctgggtgggc ccaagccacc ccagagagtt
 541   atgtttaccg aggacctgaa gctgcctgca agcttcgatg cacgggaaca atggccacag
 601   tgtcccacca tcaaagagat cagagaccag ggctcctgtg gctcctgctg ggccttcggg
 661   gctgtggaag ccatctctga ccggatctgc atccacacca atgcgcacgt cagcgtggag
 721   gtgtcggcgg aggacctgct cacatgctgt ggcagcatgt gtggggacgg ctgtaatggt
 781   ggctatcctg ctgaagcttg gaacttctgg acaagaaaag gcctggtttc tggtggcctc
 841   tatgaatccc atgtagggtg cagaccgtac tccatccctc cctgtgagca ccacgtcaac
 901   ggctcccggc ccccatgcac gggggaggga gataccccca agtgtagcaa gatctgtgag
 961   cctggctaca gcccgaccta caaacaggac aagcactacg gatacaattc ctacagcgtc
1021   tccaatagcg agaaggacat catggccgag atctacaaaa acggccccgt ggagggagct
1081   ttctctgtgt attcggactt cctgctctac aagtcaggag tgtaccaaca cgtcaccgga
1141   gagatgatgg gtggccatgc catccgcatc ctgggctggg gagtggagaa tggcacaccc
1201   tactggctgg ttgccaactc ctggaacact gactggggtg acaatggctt ctttaaaata
1261   ctcagaggac aggatcactg tggaatcgaa tcagaagtgg tggctggaat tccacgcacc
1321   gatcagtact gggaaaagat ctaatctgcc gtgggcctgt cgtgccagtc ctgggggcga
1381   gatcggggta gaaatgcatt ttattcttta agttcacgta agatacaagt ttcagacagg
1441   gtctgaagga ctggattggc caaacatcag acctgtcttc caaggagacc aagtcctggc
1501   tacatcccag cctgtggtta cagtgcagac aggccatgtg agccaccgct gccagcacag
1561   agcgtccttc cccctgtaga ctagtgccgt agggagtacc tgctgcccca gctgactgtg
1621   gccccctccg tgatccatcc atctccaggg agcaagacag agacgcagga atggaaagcg
1681   gagttcctaa caggatgaaa gttcccccat cagttccccc agtacctcca agcaagtagc
1741   tttccacatt tgtcacagaa atcagaggag agacggtgtt gggagccctt tggagaacgc
1801   cagtctccca ggccccctgc atctatcgag tttgcaatgt cacaacctct ctgatcttgt
1861   gctcagcatg attctttaat agaagtttta ttttttcgtg cactctgcta atcatgtggg
1921   tgagccagtg gaacagcggg agacctgtgc tagttttaca gattgcctcc ttatgacgcg
1981   gctcaaaagg aaaccaagtg gtcaggagtt gtttctgacc cactgatctc tactaccaca
2041   aggaaaatag tttaggagaa accagctttt actgtttttg aaaaattaca gcttcaccct
2101   gtcaagttaa caaggaatgc ctgtgccaat aaaagttttc tccaacttga agtctactct
2161   gatgggatct cagatccttt gtcactgcct atagacttgt agctgctgtc tctctttgtc
2221   cctgcagaga atcacgtcct ggaactgcat gttcttgcga ctcttgggac ttcatcttaa
2281   cttctcgctg ccccagccat gttttcaacc atggcatccc tcccccaatt agttccctgt
2341   catcctcgtc aaccttctct gtaagtgcct ggtaagcttg cccttgctta agaactcaaa
2401   acatagctgt gctctatttt tttgttgttg ttgtgactga cagagtgaga ttccgtctcc
2461   caggctggag tgcagtggcg ccttctcagc tcactgcaac ctgcagcctc ctagattcaa
2521   gcgattctcc tgcttcagcc ttccgagtag ctgggatgac aggcactcac caatatgcct
2581   gggtaatttt tgtatttttа agtacataca ggatttcacc atgttggcca ggctagtttc
2641   aaactcccgg cctcaggtgg tctgcctgcc tcagcctccc aaagtgttgg gattacaggc
2701   gtgagccact gggccctgcc tgtatttttt atcagccaca aatccagcaa caagctgagg
2761   attcagctca taaaacaggc ttggtgtctt ggtgatctca cataaccaag atgctacccc
2821   gtggggaacc acatcccccт ggatgccctc cagccttggt ttgggctgga gtcagggcct
2881   gtatacagta ttttgaattt gtatgccact ggtttgcatt gctggtcagg aactctagtg
2941   ctttgcatag ccctggttta gaaacatgtt atagcagttc ttggtataga gcaaactaga
3001   agaaccagca atcattccac tgtcctgcca aggtacacct cagtactccc cttcccaact
3061   gaagtggtat gaggctagct ctttccaaaa gcattcaagt ttggcttctg atgtgactca
3121   gaatttagga accagatgct agatcaaata agctctgaaa atctgaggaa cattgtagga
3181   aaggtttgtt aagcatctct taagtgccat gatgagcata acagccggcc gtcgtggctc
3241   acgcctgtaa tcccagcact ttgggaggcc aaggtgggag gatgacaagg tcaggagttc
3301   aagaccagcc tggccaacat gctgaaacct cacctctact aaaaatacaa aaattagctg
3361   ggcatggtgg cacatgcctg taatcccagc tacttgggag gctgaggcag gagaatcgct
3421   tgaacccggg aggcggaggt tgcagtgagc caagacagtg ccagtgcact ccagcctcgg
3481   tgacagcgca aggctccgtc tcaataatta aaaaaaaaaa aaaaaaaaaa aaggccgggc
3541   gcagtggctc aagcctgtaa tcccagcact ttgggaggct gaggcgggca gatcacctga
3601   ggtcaggagt tttgagatca gccttggcaa cacggtgaaa ccccatctct actaaaaata
3661   caaaattagc caagcatgct ggcacatgcc tgtaatccca gctactcggg aggctgaggt
3721   acgagaatcg cttgaacctg ggaggcagag gatgcagtga gccgagatca cgccattgca
3781   ctccagcctg ggggacaaga gtgaatctgt gtctcaccaa aaaaaaaaag aaaaagaaag
3841   atgcttaaca aaggttacca taagccacaa attcataacc acttatcctt ccagtttcaa
3901   gtagaatata ttcataacct caataaagtt ctccctgctc ccaaa
```

SEQ ID NO: 47 Human Cathepsin B Polypeptide, variant 2

```
MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT
YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI
```

SEQ ID NO: 48 Human Cathepsin B mRNA, variant 3

```
   1     ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgca-
       ggg
```

-continued

SEQUENCE LISTING

```
  61   ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct
 121   ctcggctgca gcgctgggtg tcttcaggcc tatggagagc agcttgcgtg ggctgggcct
 181   gcagtacctg gtttgcatag atgattggca ggtgggcagc acggggaagg acctgtgagt
 241   ggccaacctg gttcaggtgg atctaggatc cggcttccaa catgtggcag ctctgggcct
 301   ccctctgctg cctgctggtg ttggccaatg cccggagcag gccctctttc catcccctgt
 361   cggatgagct ggtcaactat gtcaacaaac ggaataccac gtggcaggcc gggcacaact
 421   tctacaacgt ggacatgagc tacttgaaga ggctatgtgg taccttcctg ggtgggccca
 481   agccacccca gagagttatg tttaccgagg acctgaagct gcctgcaagc ttcgatgcac
 541   gggaacaatg gccacagtgt cccaccatca aagagatcag agaccagggc tcctgtggct
 601   cctgctgggc cttcggggct gtggaagcca tctctgaccg gatctgcatc cacaccaatg
 661   cgcacgtcag cgtggaggtg tcggcggagg acctgctcac atgctgtggc agcatgtgtg
 721   gggacggctg taatggtggc tatcctgctg aagcttggaa cttctggaca agaaaaggcc
 781   tggtttctgg tggcctctat gaatcccatg tagggtgcag accgtactcc atccctccct
 841   gtgagcacca cgtcaacggc tcccggcccc catgcacggg ggagggagat acccccaagt
 901   gtagcaagat ctgtgagcct ggctacagcc cgacctacaa acaggacaag cactacggat
 961   acaattccta cagcgtctcc aatagcgaga aggacatcat ggccgagatc tacaaaaacg
1021   gccccgtgga gggagctttc tctgtgtatt cggacttcct gctctacaag tcaggagtgt
1081   accaacacgt caccggagag atgatgggtg gccatgccat ccgcatcctg ggctggggag
1141   tggagaatgg cacaccctac tggctggttg ccaactcctg gaacactgac tggggtgaca
1201   atggcttctt taaaatactc agaggacagg atcactgtgg aatcgaatca gaagtggtgg
1261   ctggaattcc acgcaccgat cagtactggg aaaagatcta atctgccgtg ggcctgtcgt
1321   gccagtcctg ggggcgagat cggggtagaa atgcatttta ttctttaagt tcacgtaaga
1381   tacaagtttc agacagggtc tgaaggactg gattggccaa acatcagacc tgtcttccaa
1441   ggagaccaag tcctggctac atcccagcct gtggttacag tgcagacagg ccatgtgagc
1501   caccgctgcc agcacagagc gtccttcccc ctgtagacta gtgccgtagg gagtacctgc
1561   tgccccagct gactgtggcc ccctccgtga tccatccatc tccagggagc aagacagaga
1621   cgcaggaatg gaaagcggag ttcctaacag gatgaaagtt cccccatcag ttcccccagt
1681   acctccaagc aagtagcttt ccacatttgt cacagaaatc agaggagaga cggtgttggg
1741   agccctttgg agaacgccag tctcccaggc cccctgcatc tatcgagttt gcaatgtcac
1801   aacctctctg atcttgtgct cagcatgatt ctttaataga agttttattt tttcgtgcac
1861   tctgctaatc atgtgggtga gccagtggaa cagcgggaga cctgtgctag ttttacagat
1921   tgcctcctta tgacgcggct caaaaggaaa ccaagtggtc aggagttgtt tctgacccac
1981   tgatctctac taccacaagg aaaatagttt aggagaaacc agcttttact gtttttgaaa
2041   aattacagct tcaccctgtc aagttaacaa ggaatgcctg tgccaataaa agtttttctcc
2101   aacttgaagt ctactctgat gggatctcag atcctttgtc actgcctata gacttgtagc
2161   tgctgtctct ctttgtccct gcagagaatc acgtcctgga actgcatgtt cttgcgactc
2221   ttgggacttc atcttaactt ctcgctgccc cagccatgtt ttcaaccatg gcatccctcc
2281   cccaattagt tccctgtcat cctcgtcaac cttctctgta agtgcctggt aagcttgccc
2341   ttgcttaaga actcaaaaca tagctgtgct ctatttttt gttgttgttg tgactgacag
2401   agtgagattc cgtctcccag gctggagtgc agtggcgcct tctcagctca ctgcaacctg
2461   cagcctccta gattcaagcg attctcctgc ttcagccttc cgagtagctg ggatgacagg
2521   cactcaccaa tatgcctggg taatttttgt atttttaagt acatacagga tttcaccatg
2581   ttggccaggc tagtttcaaa ctcccggcct caggtggtct gcctgcctca gcctcccaaa
2641   gtgttgggat tacaggcgtg agccactggg ccctgcctgt atttttatc agccacaaat
2701   ccagcaacaa gctgaggatt cagctcataa aacaggcttg gtgtcttggt gatctcacat
2761   aaccaagatg ctaccccgtg gggaaccaca tccccctgga tgccctccag ccttggtttg
2821   ggctggagtc agggcctgta tacagtattt tgaatttgta tgccactggt ttgcattgct
2881   ggtcaggaac tctagtgctt tgcatagccc tggtttagaa acatgttata gcagttcttg
2941   gtatagagca aactagaaga accagcaatc attccactgt cctgccaagg tacacctcag
3001   tactcccctt cccaactgaa gtggtatgag gctagctctt tccaaaagca ttcaagtttg
3061   gcttctgatg tgactcagaa tttaggaacc agatgctaga tcaaataagc tctgaaaatc
3121   tgaggaacat tgtaggaaag gtttgttaag catctcttaa gtgccatgat gagcataaca
3181   gccggccgtc gtggctcacg cctgtaatcc cagcactttg ggaggccaag gtgggaggat
3241   gacaaggtca ggagttcaag accagcctgg ccaacatgct gaaacctcac ctctactaaa
3301   aatacaaaaa ttagctgggc atggtggcac atgcctgtaa tcccagctac ttgggaggct
3361   gaggcaggag aatcgcttga acccgggagg cggaggttgc agtgagccaa gacagtgcca
3421   gtgcactcca gcctcggtga cagcgcaagg ctccgtctca ataattaaaa aaaaaaaaaa
3481   aaaaaaaaag gccgggcgca gtggctcaag cctgtaatcc cagcactttg ggaggctgag
3541   gcgggcagat cacctgaggt caggagtttt gagatcagcc ttggcaacac ggtgaaaccc
3601   catctctact aaaaatacaa aattagccaa gcatgctggc acatgcctgt aatcccagct
3661   actcgggagg ctgaggtacg agaatcgctt gaacctggga ggcagaggat gcagtgagcc
3721   gagatcacgc cattgcactc cagcctgggg gacaagagtg aatctgtgtc tcaccaaaaa
3781   aaaaaagaaa aagaaagatg cttaacaaag gttaccataa gccacaaatt cataaccact
3841   tatccttcca gtttcaagta gaatatattc ataacctcaa taaagttctc cctgctccca
3901   aa
```

SEQ ID NO: 49 Human Cathepsin B Polypeptide, variant 3

```
MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT
YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI
```

-continued

SEQUENCE LISTING

SEQ ID NO: 50 Human Cathepsin B mRNA, variant 4

```
   1   ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg
  61   ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct
 121   ctcggctgca gcgctgggct ggtgtgcagt ggtgcgacca cggctcacgg cagcctcagc
 181   cacccagatg taagcgatct ggttcccacc tcagcctccc gagtagtgga tctaggatcc
 241   ggcttccaac atgtggcagc tctgggcctc cctctgctgc ctgctggtgt tggccaatgc
 301   ccggagcagg ccctctttcc atcccctgtc ggatgagctg gtcaactatg tcaacaaacg
 361   gaataccacg tggcaggccg ggcacaactt ctacaacgtg gacatgagct acttgaagag
 421   gctatgtggt accttcctgg gtgggcccaa gccaccccag agagttatgt ttaccgagga
 481   cctgaagctg cctgcaagct tcgatgcacg ggaacaatgg ccacagtgtc ccaccatcaa
 541   agagatcaga gaccagggct cctgtggctc ctgctgggcc ttcggggctg tggaagccat
 601   ctctgaccgg atctgcatcc acaccaatgc gcacgtcagc gtggaggtgt cggcggagga
 661   cctgctcaca tgctgtggca gcatgtgtgg ggacggctgt aatggtggct atcctgctga
 721   agcttggaac ttctggacaa gaaaaggcct ggtttctggt ggcctctatg aatcccatgt
 781   agggtgcaga ccgtactcca tcctccctg tgagcaccac gtcaacggct cccggccccc
 841   atgcacgggg gagggagata cccccaagtg tagcaagatc tgtgagcctg gctacagccc
 901   gacctacaaa caggacaagc actacggata caattcctac agcgtctcca atagcgagaa
 961   ggacatcatg gccgagatct acaaaaacgg ccccgtggag ggagctttct ctgtgtattc
1021   ggacttcctg ctctacaagt caggagtgta ccaacacgtc accggagaga tgatgggtgg
1081   ccatgccatc cgcatcctgg gctggggagt ggagaatggc acaccctact ggctggttgc
1141   caactcctgg aacactgact ggggtgacaa tggcttcttt aaaatactca gaggacagga
1201   tcactgtgga atcgaatcag aagtggtggc tggaattcca cgcaccgatc agtactggga
1261   aaagatctaa tctgccgtgg gcctgtcgtg ccagtcctgg gggcgagatc ggggtagaaa
1321   tgcattttat tctttaagtt cacgtaagat acaagtttca gacagggtct gaaggactgg
1381   attggccaaa catcagacct gtcttccaag gagaccaagt cctggctaca tcccagcctg
1441   tggttacagt gcagacaggc catgtgagcc accgctgcca gcacagagcg tccttccccc
1501   tgtagactag tgccgtaggg agtacctgct gccccagctg actgtggccc cctccgtgat
1561   ccatccatct ccagggagca agacagagac gcaggaatgg aaagcggagt tcctaacagg
1621   atgaaagttc ccccatcagt tcccccagta cctccaagca agtagctttc cacatttgtc
1681   acagaaatca gaggagagac ggtgttggga gccctttgga gaacgccagt ctcccaggcc
1741   ccctgcatct atcgagtttg caatgtcaca acctctctga tcttgtgctc agcatgattc
1801   tttaatagaa gttttatttt ttcgtgcact ctgctaatca tgtgggtgag ccagtggaac
1861   agcgggagac ctgtgctagt tttacagatt gcctccttat gacgcggctc aaaaggaaac
1921   caagtggtca ggagttgttt ctgacccact gatctctact accacaagga aaatagttta
1981   ggagaaacca gcttttactg tttttgaaaa attacagctt caccctgtca agttaacaag
2041   gaatgcctgt gccaataaaa gttttctcca acttgaagtc tactctgatg ggatctcaga
2101   tcctttgtca ctgcctatag acttgtagct gctgtctctc tttgtccctg cagagaatca
2161   cgtcctggaa ctgcatgttc ttgcgactct tgggacttca tcttaacttc tcgctgcccc
2221   agccatgttt tcaaccatgg catccctccc ccaattagtt ccctgtcatc ctcgtcaacc
2281   ttctctgtaa gtgcctggta agcttgccct tgcttaagaa ctcaaaacat agctgtgctc
2341   tatttttttg ttgttgttgt gactgacaga gtgagattcc gtctcccagg ctggagtgca
2401   gtggcgcctt ctcagctcac tgcaacctgc agcctcctag attcaagcga ttctcctgct
2461   tcagccttcc gagtagctgg gatgacaggc actcaccaat atgcctgggt aatttttgta
2521   tttttaagta catacaggat ttcaccatgt tggccaggct agtttcaaac tcccggcctc
2581   aggtggtctg cctgcctcag cctcccaaag tgttgggatt acaggcgtga gccactgggc
2641   cctgcctgta ttttttatca gccacaaatc cagcaacaag ctgaggattc agctcataaa
2701   acaggcttgg tgtcttggtg atctcacata accaagatgc taccccgtgg ggaaccacat
2761   ccccctggat gccctccagc cttggtttgg gctggagtca gggcctgtat acagtatttt
2821   gaatttgtat gccactggtt tgcattgctg gtcaggaact ctagtgcttt gcatagccct
2881   ggtttagaaa catgttatag cagttcttgg tatagagcaa actagaagaa ccagcaatca
2941   ttccactgtc ctgccaaggt acacctcagt actccccttc ccaactgaag tggtatgagg
3001   ctagctcttt ccaaaagcat tcaagtttgg cttctgatgt gactcagaat ttaggaacca
3061   gatgctagat caaataagct ctgaaaatct gaggaacatt gtaggaaagg tttgttaagc
3121   atctcttaag tgccatgatg agcataacag ccggccgtcg tggctcacgc ctgtaatccc
3181   agcactttgg gaggccaagg tgggaggatg acaaggtcag gagttcaaga ccagcctggc
3241   caacatgctg aaacctcacc tctactaaaa atacaaaaat tagctgggca tggtggcaca
3301   tgcctgtaat cccagctact tgggaggctg aggcaggaga atcgcttgaa cccgggaggc
3361   ggaggttgca gtgagccaag acagtgccag tgcactccag cctcggtgac agcgcaaggc
3421   tccgtctcaa taattaaaaa aaaaaaaaaa aaaaaaaagg ccgggcgcag tggctcaagc
3481   ctgtaatccc agcactttgg gaggctgagg cgggcagatc acctgaggtc aggagttttg
3541   agatcagcct tggcaacacg gtgaaacccc atctctacta aaaatacaaa attagccaag
3601   catgctggca catgcctgta atcccagcta ctcgggaggc tgaggtacga gaatcgcttg
3661   aacctgggag gcagaggatg cagtgagccg agatcacgcc attgcactcc agcctggggg
3721   acaagagtga atctgtgtct caccaaaaaa aaaaagaaaa agaaagatgc ttaacaaagg
3781   ttaccataag ccacaaattc ataaccactt atccttccag tttcaagtag aatatattca
3841   taacctcaat aaagttctcc ctgctcccaa a
```

SEQ ID NO: 51 Human Cathepsin B Polypeptide, variant 4

```
MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT
YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
```

-continued

SEQUENCE LISTING

GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI

SEQ ID NO: 52 Human Cathepsin B mRNA, variant 5

```
   1   ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg
  61   ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct
 121   ctcggctgca gcgctgggtg tcttcaggcc tatggagagc agcttgcgtg ggctgggcct
 181   gcagtacctg gtttgcatag atgattggca ggtggatcta ggatccggct tccaacatgt
 241   ggcagctctg ggcctccctc tgctgcctgc tggtgttggc caatgcccgg agcaggccct
 301   ctttccatcc cctgtcggat gagctggtca actatgtcaa caaacggaat accacgtggc
 361   aggccgggca caacttctac aacgtggaca tgagctactt gaagaggcta tgtggtacct
 421   tcctgggtgg gcccaagcca ccccagagag ttatgtttac cgaggacctg aagctgcctg
 481   caagcttcga tgcacgggaa caatggccac agtgtcccac catcaaagag atcagagacc
 541   agggctcctg tggctcctgc tgggccttcg gggctgtgga agccatctct gaccggatct
 601   gcatccacac caatgcgcac gtcagcgtgg aggtgtcggc ggaggacctg ctcacatgct
 661   gtggcagcat gtgtggggac ggctgtaatg gtggctatcc tgctgaagct tggaacttct
 721   ggacaagaaa aggcctggtt tctggtggcc tctatgaatc ccatgtaggg tgcagaccgt
 781   actccatccc tccctgtgag caccacgtca acggctcccg gcccccatgc acgggggagg
 841   gagatacccc caagtgtagc aagatctgtg agcctggcta cagcccgacc tacaaacagg
 901   acaagcacta cggatacaat tcctacagcg tctccaatag cgagaaggac atcatggccg
 961   agatctacaa aaacggcccc gtggagggag ctttctctgt gtattcggac ttcctgctct
1021   acaagtcagg agtgtaccaa cacgtcaccg gagagatgat gggtggccat gccatccgca
1081   tcctgggctg gggagtggag aatggcacac cctactggct ggttgccaac tcctggaaca
1141   ctgactgggg tgacaatggc ttctttaaaa tactcagagg acaggatcac tgtggaatcg
1201   aatcagaagt ggtggctgga attccacgca ccgatcagta ctgggaaaag atctaatctg
1261   ccgtgggcct gtcgtgccag tcctgggggc gagatcgggg tagaaatgca ttttattctt
1321   taagttcacg taagatacaa gtttcagaca gggtctgaag gactggattg gccaaacatc
1381   agacctgtct tccaaggaga ccaagtcctg gctacatccc agcctgtggt tacagtgcag
1441   acaggccatg tgagccaccg ctgccagcac agagcgtcct tccccctgta gactagtgcc
1501   gtagggagta cctgctgccc cagctgactg tggccccctc cgtgatccat ccatctccag
1561   ggagcaagac agagacgcag gaatggaaag cggagttcct aacaggatga aagttccccc
1621   atcagttccc ccagtacctc caagcaagta gctttccaca tttgtcacag aaatcagagg
1681   agagacggtg ttgggagccc tttggagaac gccagtctcc caggcccctc gcatctatcg
1741   agtttgcaat gtcacaacct ctctgatctt gtgctcagca tgattcttta atagaagttt
1801   tatttttttcg tgcactctgc taatcatgtg ggtgagccag tggaacagcg ggagacctgt
1861   gctagttttta cagattgcct ccttatgacg cggctcaaaa ggaaaccaag tggtcaggag
1921   ttgtttctga cccactgatc tctactacca caaggaaaat agtttaggag aaaccagctt
1981   ttactgtttt tgaaaaatta cagcttcacc ctgtcaagtt aacaaggaat gcctgtgcca
2041   ataaaagttt tctccaactt gaagtctact ctgatgggat ctcagatcct ttgtcactgc
2101   ctatagactt gtagctgctg tctctctttg tccctgcaga gaatcacgtc ctggaactgc
2161   atgttcttgc gactcttggg acttcatctt aacttctcgc tgccccagcc atgttttcaa
2221   ccatggcatc cctcccccaa ttagttccct gtcatcctcg tcaaccttct ctgtaagtgc
2281   ctggtaagct tgcccttgct taagaactca aaacatagct gtgctctatt tttttgttgt
2341   tgttgtgact gacagagtga gattccgtct cccaggctgg agtgcagtgg cgccttctca
2401   gctcactgca acctgcagcc tcctagattc aagcgattct cctgcttcag ccttccgagt
2461   agctgggatg acaggcactc accaatatgc ctgggtaatt tttgtatttt taagtacata
2521   caggatttca ccatgttggc caggctagtt tcaaactccc ggcctcaggt ggtctgcctg
2581   cctcagcctc ccaaagtgtt gggattacag gcgtgagcca ctgggccctg cctgtatttt
2641   ttatcagcca caaatccagc aacaagctga ggattcagct cataaaacag gcttggtgtc
2701   ttggtgatct cacataacca agatgctacc ccgtggggaa ccacatcccc ctggatgccc
2761   tccagccttg gtttgggctg gagtcagggc ctgtatacag tattttgaat ttgtatgcca
2821   ctggtttgca ttgctggtca ggaactctag tgctttgcat agccctggtt tagaaacatg
2881   ttatagcagt tcttggtata gagcaaacta gaagaaccag caatcattcc actgtcctgc
2941   caaggtacac ctcagtactc cccttcccaa ctgaagtggt atgaggctag ctctttccaa
3001   aagcattcaa gtttggcttc tgatgtgact cagaatttag gaaccagatg ctagatcaaa
3061   taagctctga aaatctgagg aacattgtag gaaaggtttg ttaagcatct cttaagtgcc
3121   atgatgagca taacagccgg ccgtcgtggc tcacgcctgt aatcccagca ctttgggagg
3181   ccaaggtggg aggatgacaa ggtcaggagt tcaagaccag cctggccaac atgctgaaac
3241   ctcacctcta ctaaaaatac aaaaattagc tgggcatggt ggcacatgcc tgtaatccca
3301   gctacttggg aggctgaggc aggagaatcg cttgaacccg ggaggcggag gttgcagtga
3361   gccaagacag tgccagtgca ctccagcctc ggtgacagcg caaggctccg tctcaataat
3421   taaaaaaaaa aaaaaaaaaa aaaaggccgg gcgcagtggc tcaagcctgt aatcccagca
3481   ctttgggagg ctgaggcggg cagatcacct gaggtcagga gttttgagat cagccttggc
3541   aacacggtga aaccccatct ctactaaaaa tacaaaatta gccaagcatg ctggcacatg
3601   cctgtaatcc cagctactcg ggaggctgag gtacgagaat cgcttgaacc tgggaggcag
3661   aggatgcagt gagccgagat cacgccattg cactccagcc tgggggacaa gagtgaatct
3721   gtgtctcacc aaaaaaaaaa agaaaaagaa agatgcttaa caaaggttac cataagccac
3781   aaattcataa ccacttatcc ttccagtttc aagtagaata tattcataac ctcaataaag
3841   ttctccctgc tcccaaa
```

SEQ ID NO: 53 Human Cathepsin B Polypeptide, variant 5

MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT

-continued

SEQUENCE LISTING

YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI

SEQ ID NO: 54 Human Cathepsin B mRNA, variant 6

```
   1   agggccgggg ctggcccagg ctacggcggc tgcagggctc cggcaaccgc tccggcaacg
  61   ccaaccgctc cgctgcgcgc aggctgggct gcaggctctc ggctgcagcg ctgggctggt
 121   gtgcagtggt gcgaccacgg ctcacggcag cctcagccac ccagatgtaa gcgatctggt
 181   tcccacctca gcctcccgag tagatacttc tgaaaatага aatgatgact ctgggatgca
 241   aacgttggct gtcctatgta taaggagatg gcttttcacg ctcccagtga ctgaggaagt
 301   ttctcccaga tggcgctgct ctgagcctgg tgcagggtgg atctaggatc cggcttccaa
 361   catgtggcag ctctgggcct ccctctgctg cctgctggtg ttggccaatg cccggagcag
 421   gccctctttc catcccctgt cggatgagct ggtcaactat gtcaacaaac ggaataccac
 481   gtggcaggcc gggcacaact tctacaacgt ggacatgagc tacttgaaga ggctatgtgg
 541   taccttcctg ggtgggccca agccacccca gagagttatg tttaccgagg acctgaagct
 601   gcctgcaagc ttcgatgcac gggaacaatg gccacagtgt cccaccatca aagagatcag
 661   agaccagggc tcctgtggct cctgctgggc cttcggggct gtggaagcca tctctgaccg
 721   gatctgcatc cacaccaatg cgcacgtcag cgtggaggtg tcggcggagg acctgctcac
 781   atgctgtggc agcatgtgtg gggacggctg taatggtggc tatcctgctg aagcttggaa
 841   cttctggaca agaaaaggcc tggtttctgg tggcctctat gaatcccatg tagggtgcag
 901   accgtactcc atccctccct gtgagcacca cgtcaacggc tcccggcccc catgcacggg
 961   ggagggagat acccccaagt gtagcaagat ctgtgagcct ggctacagcc cgacctacaa
1021   acaggacaag cactacggat acaattccta cagcgtctcc aatagcgaga aggacatcat
1081   ggccgagatc tacaaaaacg gccccgtgga gggagctttc tctgtgtatt cggacttcct
1141   gctctacaag tcaggagtgt accaacacgt caccggagag atgatgggtg gccatgccat
1201   ccgcatcctg ggctggggag tggagaatgg cacaccctac tggctggttg ccaactcctg
1261   gaacactgac tggggtgaca atggcttctt taaaatactc agaggacagg atcactgtgg
1321   aatcgaatca gaagtggtgg ctggaattcc acgcaccgat cagtactggg aaaagatcta
1381   atctgccgtg ggcctgtcgt gccagtcctg ggggcgagat cggggtagaa atgcatttta
1441   ttctttaagt tcacgtaaga tacaagtttc agacagggtc tgaaggactg gattggccaa
1501   acatcagacc tgtcttccaa ggagaccaag tcctggctac atcccagcct gtggttacag
1561   tgcagacagg ccatgtgagc caccgctgcc agcacagagc gtccttcccc ctgtagacta
1621   gtgccgtagg gagtacctgc tgccccagct gactgtggcc ccctccgtga tccatccatc
1681   tccagggagc aagacagaga cgcaggaatg gaaagcggag ttcctaacag gatgaaagtt
1741   cccccatcag ttcccccagt acctccaagc aagtagcttt ccacatttgt cacagaaatc
1801   agaggagaga cggtgttggg agccctttgg agaacgccag tctcccaggc cccctgcatc
1861   tatcgagttt gcaatgtcac aacctctctg atcttgtgct cagcatgatt ctttaataga
1921   agttttattt tttcgtgcac tctgctaatc atgtgggtga gccagtggaa cagcgggaga
1981   cctgtgctag ttttacagat tgcctcctta tgacgcggct caaaaggaaa ccaagtggtc
2041   aggagttgtt tctgacccac tgatctctac taccacaagg aaaatagttt aggagaaacc
2101   agcttttact gtttttgaaa aattacagct tcaccctgtc aagttaacaa ggaatgcctg
2161   tgccaataaa agttttctcc aacttgaagt ctactctgat gggatctcag atcctttgtc
2221   actgcctata gacttgtagc tgctgtctct ctttgtccct gcagagaatc acgtcctgga
2281   actgcatgtt cttgcgactc ttgggacttc atcttaactt ctcgctgccc cagccatgtt
2341   ttcaaccatg gcatccctcc cccaattagt tccctgtcat cctcgtcaac cttctctgta
2401   agtgcctggt aagcttgccc ttgcttaaga actcaaaaca tagctgtgct ctatttttt
2461   gttgttgttg tgactgacag agtgagattc cgtctcccag gctggagtgc agtggcgcct
2521   tctcagctca ctgcaacctg cagcctccta gattcaagcg attctcctgc ttcagccttc
2581   cgagtagctg ggatgacagg cactcaccaa tatgcctggg taatttttgt atttttaagt
2641   acatacagga tttcaccatg ttggccaggc tagtttcaaa ctcccggcct caggtggtct
2701   gcctgcctca gcctcccaaa gtgttgggat tacaggcgtg agccactggg ccctgcctgt
2761   atttttatc agccacaaat ccagcaacaa gctgaggatt cagctcataa aacaggcttg
2821   gtgtcttggt gatctcacat aaccaagatg ctaccccgtg gggaaccaca tccccctgga
2881   tgccctccag ccttggtttg ggctggagtc agggcctgta tacagtattt tgaatttgta
2941   tgccactggt ttgcattgct ggtcaggaac tctagtgctt tgcatagccc tggtttagaa
3001   acatgttata gcagttcttg gtatagagca aactagaaga accagcaatc attccactgt
3061   cctgccaagg tacacctcag tactcccctt cccaactgaa gtggtatgag gctagctctt
3121   tccaaaagca ttcaagtttg gcttctgatg tgactcagaa tttaggaacc agatgctaga
3181   tcaaataagc tctgaaaatc tgaggaacat tgtaggaaag gtttgttaag catctcttaa
3241   gtgccatgat gagcataaca gccggccgtc gtggctcacg cctgtaatcc cagcactttg
3301   ggaggccaag gtgggaggat gacaaggtca ggagttcaag accagcctgg ccaacatgct
3361   gaaacctcac ctctactaaa aatacaaaaa ttagctgggc atggtggcac atgcctgtaa
3421   tcccagctac ttgggaggct gaggcaggag aatcgcttga acccgggagg cggaggttgc
3481   agtgagccaa gacagtgcca gtgcactcca gcctcggtga cagcgcaagg ctccgtctca
3541   ataattaaaa aaaaaaaaaa aaaaaaaaag gccgggcgca gtggctcaag cctgtaatcc
3601   cagcactttg ggaggctgag gcgggcagat cacctgaggt caggagtttt gagatcagcc
3661   ttggcaacac ggtgaaaccc catctctact aaaaatacaa aattagccaa gcatgctggc
3721   acatgcctgt aatcccagct actcgggagg ctgaggtacg agaatcgctt gaacctggga
3781   ggcagaggat gcagtgagcc gagatcacgc cattgcactc cagcctgggg gacaagagtg
3841   aatctgtgtc tcaccaaaaa aaaaaagaaa aagaaagatg cttaacaaag gttaccataa
3901   gccacaaatt cataaccact tatccttcca gtttcaagta gaatatattc ataacctcaa
3961   taaagttctc cctgctccca aa
```

-continued

SEQUENCE LISTING

SEQ ID NO: 55 Human Cathepsin B Polypeptide, variant 6

MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT
YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI

SEQ ID NO: 56 Human Cathepsin B mRNA, variant 7

```
   1   caggaccgcc gagggaggcg cctgcgagga agagctcggc cgggtccgga gactgctgcc
  61   tgggaccgcg ctcccagcgc ctgggcctcg gtgtctccgg gccaaactgc cgacataatc
 121   gcatctgccg gcatctattt tcggtttatt tccccctcat tgcgaaggat ttgcctggcc
 181   aactttctgc gcaagatccc acgcaattcc tgggacccca gaagacaggt cctgttgaag
 241   aacaggaatc tggcactggg tgggctgggg aggaagccgc acggtgttaa atccataaac
 301   aggaagagaa accagacagc gaaaccaaga ggcgaatggg cgattggatg ccggtgggga
 361   gaaggccggg ggcgcaccct gctcctggac tccagtaaag ggaggccggg cagagtccct
 421   ggggcgccac ctccccctcg gtggatctag gatccggctt ccaacatgtg gcagctctgg
 481   gcctccctct gctgcctgct ggtgttggcc aatgcccgga gcaggccctc tttccatccc
 541   ctgtcggatg agctggtcaa ctatgtcaac aaacggaata ccacgtggca ggccgggcac
 601   aacttctaca acgtggacat gagctacttg aagaggctat gtggtacctt cctgggtggg
 661   cccaagccac cccagagagt tatgtttacc gaggacctga agctgcctgc aagcttcgat
 721   gcacgggaac aatggccaca gtgtcccacc atcaaagaga tcagagacca gggctcctgt
 781   ggctcctgct gggccttcgg ggctgtggaa gccatctctg accggatctg catccacacc
 841   aatgcgcacg tcagcgtgga ggtgtcggcg gaggacctgc tcacatgctg tggcagcatg
 901   tgtggggacg gctgtaatgg tggctatcct gctgaagctt ggaacttctg gacaagaaaa
 961   ggcctggttt ctggtggcct ctatgaatcc catgtagggt gcagaccgta ctccatccct
1021   ccctgtgagc accacgtcaa cggctcccgg cccccatgca cgggggaggg agataccccc
1081   aagtgtagca agatctgtga gcctggctac agcccgacct acaaacagga caagcactac
1141   ggatacaatt cctacagcgt ctccaatagc gagaaggaca tcatggccga gatctacaaa
1201   aacggccccg tggagggagc tttctctgtg tattcggact tcctgctcta caagtcagga
1261   gtgtaccaac acgtcaccgg agagatgatg ggtggccatg ccatccgcat cctgggctgg
1321   ggagtggaga atggcacacc ctactggctg gttgccaact cctggaacac tgactggggt
1381   gacaatggct tctttaaaat actcagagga caggatcact gtggaatcga atcagaagtg
1441   gtggctggaa ttccacgcac cgatcagtac tgggaaaaga tctaatctgc cgtgggcctg
1501   tcgtgccagt cctgggggcg agatcggggt agaaatgcat tttattcttt aagttcacgt
1561   aagatacaag tttcagacag ggtctgaagg actggattgg ccaaacatca gacctgtctt
1621   ccaaggagac caagtcctgg ctacatccca gcctgtggtt acagtgcaga caggccatgt
1681   gagccaccgc tgccagcaca gagcgtcctt ccccctgtag actagtgccg tagggagtac
1741   ctgctgcccc agctgactgt ggccccctcc gtgatccatc catctccagg gagcaagaca
1801   gagacgcagg aatggaaagc ggagttccta acaggatgaa agttccccca tcagttcccc
1861   cagtacctcc aagcaagtag ctttccacat ttgtcacaga aatcagagga gagacggtgt
1921   tgggagccct ttggagaacg ccagtctccc aggccccctg catctatcga gtttgcaatg
1981   tcacaacctc tctgatcttg tgctcagcat gattctttaa tagaagtttt atttttcgt
2041   gcactctgct aatcatgtgg gtgagccagt ggaacagcgg gagacctgtg ctagttttac
2101   agattgcctc cttatgacgc ggctcaaaag gaaaccaagt ggtcaggagt tgtttctgac
2161   ccactgatct ctactaccac aaggaaaata gtttaggaga aaccagcttt tactgttttt
2221   gaaaaattac agcttcaccc tgtcaagtta acaaggaatg cctgtgccaa taaaagtttt
2281   ctccaacttg aagtctactc tgatgggatc tcagatcctt tgtcactgcc tatagacttg
2341   tagctgctgt ctctctttgt ccctgcagag aatcacgtcc tggaactgca tgttcttgcg
2401   actcttggga cttcatctta acttctcgct gccccagcca tgttttcaac catggcatcc
2461   ctcccccaat tagttccctg tcatcctcgt caaccttctc tgtaagtgcc tggtaagctt
2521   gcccttgctt aagaactcaa aacatagctg tgctctattt ttttgttgtt gttgtgactg
2581   acagagtgag attccgtctc ccaggctgga gtgcagtggc gccttctcag ctcactgcaa
2641   cctgcagcct cctagattca agcgattctc ctgcttcagc cttccgagta gctgggatga
2701   caggcactca ccaatatgcc tgggtaattt ttgtattttt aagtacatac aggatttcac
2761   catgttggcc aggctagttt caaactcccg gcctcaggtg gtctgcctgc ctcagcctcc
2821   caaagtgttg ggattacagg cgtgagccac tgggccctgc ctgtatttt tatcagccac
2881   aaatccagca acaagctgag gattcagctc ataaaacagg cttggtgtct tggtgatctc
2941   acataaccaa gatgctaccc cgtggggaac cacatccccc tggatgccct ccagccttgg
3001   tttgggctgg agtcagggcc tgtatacagt attttgaatt tgtatgccac tggtttgcat
3061   tgctggtcag gaactctagt gctttgcata gccctggttt agaaacatgt tatagcagtt
3121   cttggtatag agcaaactag aagaaccagc aatcattcca ctgtcctgcc aaggtacacc
3181   tcagtactcc ccttcccaac tgaagtggta tgaggctagc tctttccaaa agcattcaag
3241   tttggcttct gatgtgactc agaatttagg aaccagatgc tagatcaaat aagctctgaa
3301   aatctgagga acattgtagg aaaggtttgt taagcatctc ttaagtgcca tgatgagcat
3361   aacagccggc cgtcgtggct cacgcctgta atcccagcac tttgggaggc caaggtggga
3421   ggatgacaag gtcaggagtt caagaccagc ctggccaaca tgctgaaacc tcacctctac
3481   taaaaataca aaaattagct gggcatggtg gcacatgcct gtaatcccag ctacttggga
3541   ggctgaggca ggagaatcgc ttgaacccgg gaggcggagg ttgcagtgag ccaagacagt
3601   gccagtgcac tccagcctcg gtgacagcgc aaggctccgt ctcaataatt aaaaaaaaaa
3661   aaaaaaaaaa aaaggccggg cgcagtggct caagcctgta atcccagcac tttgggaggc
3721   tgaggcgggc agatcacctg aggtcaggag ttttgagatc agccttggca acacggtgaa
3781   accccatctc tactaaaaat acaaaattag ccaagcatgc tggcacatgc ctgtaatccc
3841   agctactcgg gaggctgagg tacgagaatc gcttgaacct gggaggcaga ggatgcagtg
```

-continued

SEQUENCE LISTING

```
3901   agccgagatc acgccattgc actccagcct gggggacaag agtgaatctg tgtctcacca
3961   aaaaaaaaaa gaaaaagaaa gatgcttaac aaaggttacc ataagccaca aattcataac
4021   cacttatcct tccagtttca agtagaatat attcataacc tcaataaagt tctccctgct
4081   cccaaa

SEQ ID NO: 57 Human Cathepsin B Polypeptide, variant 7
MWQLWASLCCLLVLANARSRPSFHPLSDELVNYVNKRNTTWQAG
HNFYNVDMSYLKRLCGTFLGGPKPPQRVMFTEDLKLPASFDAREQWPQCPTIKEIRDQ
GSCGSCWAFGAVEAISDRICIHTNAHVSVEVSAEDLLICCGSMCGDGCNGGYPAEAWN
FWIRKGLVSGGLYESHVGCRPYSIPPCEHHVNGSRPPCTGEGDTPKCSKICEPGYSPT
YKQDKHYGYNSYSVSNSEKDIMAEIYKNGPVEGAFSVYSDFLLYKSGVYQHVTGEMMG
GHAIRILGWGVENGTPYWLVANSWNIDWGDNGFFKILRGQDHCGIESEVVAGIPRIDQ
YWEKI
SEQ ID NO: 58 Human Cathepsin L mRNA, variant 2
   1   ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc
  61   agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac
 121   cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag
 181   gacgcggtcg agtaggtttt aaaacatgaa tcctacactc atccttgctg ccttttgcct
 241   gggaattgcc tcagctactc taacatttga tcacagttta gaggcacagt ggaccaagtg
 301   gaaggcgatg cacaacagat tatacggcat gaatgaagaa ggatggagga gagcagtgtg
 361   ggagaagaac atgaagatga ttgaactgca caatcaggaa tacagggaag ggaaacacag
 421   cttcacaatg gccatgaacg cctttggaga catgaccagt gaagaattca ggcaggtgat
 481   gaatggcttt caaaaccgta agcccaggaa ggggaaagtg ttccaggaac ctctgtttta
 541   tgaggccccc agatctgtgg attggagaga gaaaggctac gtgactcctg tgaagaatca
 601   gggtcagtgt ggttcttgtt gggcttttag tgctactggt gctcttgaag gacagatgtt
 661   ccggaaaact gggaggctta tctcactgag tgagcagaat ctggtagact gctctgggcc
 721   tcaaggcaat gaaggctgca atggtggcct aatggattat gctttccagt atgttcagga
 781   taatggaggc ctggactctg aggaatccta tccatatgag gcaacagaag aatcctgtaa
 841   gtacaatccc aagtattctg ttgctaatga caccggcttt gtggacatcc ctaagcagga
 901   gaaggccctg atgaaggcag ttgcaactgt ggggcccatt tctgttgcta ttgatgcagg
 961   tcatgagtcc ttcctgttct ataaagaagg catttatttt gagccagact gtagcagtga
1021   agacatggat catggtgtgc tggtggttgg ctacggattt gaaagcacag aatcagataa
1081   caataaatat tggctggtga agaacagctg gggtgaagaa tggggcatgg gtggctacgt
1141   aaagatggcc aaagaccgga gaaaccattg tggaattgcc tcagcagcca gctaccccac
1201   tgtgtgagct ggtggacggt gatgaggaag gacttgactg gggatggcgc atgcatggga
1261   ggaattcatc ttcagtctac cagcccccgc tgtgtcggat acacactcga atcattgaag
1321   atccgagtgt gatttgaatt ctgtgatatt ttcacactgg taaatgttac ctctatttta
1381   attactgcta taaataggtt tatattattg attcacttac tgactttgca ttttcgtttt
1441   taaaaggatg tataaatttt tacctgttta aataaaattt aatttcaaat gtagtggtgg
1501   ggcttctttc tatttttgat gcactgaatt tttgtgtaat aaagaacata attgggctct
1561   aagccataaa aaaaaaaaaa aaaaaaa

SEQ ID NO: 59 Human Cathepsin L Polypeptide, variant 2
MNPTLILAAFCLGIASATLTFDHSLEAQWTKWKAMHNRLYGMNE
EGWRRAVWEKNMKMIELHNQEYREGKHSFTMAMNAFGDMTSEEFRQVMNGFQNRKPRK
GKVFQEPLFYEAPRSVDWREKGYVTPVKNQGQCGSCWAFSATGALEGQMFRKTGRLIS
LSEQNLVDCSGPQGNEGCNGGLMDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYS
VANDTGFVDIPKQEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDH
GVLVVGYGFESTESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV

SEQ ID NO: 60 Human Cathepsin L mRNA, variant 3
   1   ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc
  61   agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac
 121   cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag
 181   gacgcggtcg agtaggtgtg caccagccct ggcaacgaga gcgtctaccc cgaactctgc
 241   tggccttgag gttttaaaac atgaatccta cactcatcct tgctgccttt tgcctgggaa
 301   ttgcctcagc tactctaaca tttgatcaca gtttagaggc acagtggacc aagtggaagg
 361   cgatgcacaa cagattatac ggcatgaatg aagaaggatg gaggagagca gtgtgggaga
 421   agaacatgaa gatgattgaa ctgcacaatc aggaatacag ggaagggaaa cacagcttca
 481   caatggccat gaacgccttt ggagacatga ccagtgaaga attcaggcag gtgatgaatg
 541   gctttcaaaa ccgtaagccc aggaagggga aagtgttcca ggaacctctg ttttatgagg
 601   cccccagatc tgtggattgg agagagaaag gctacgtgac tcctgtgaag aatcagggtc
 661   agtgtggttc ttgttgggct tttagtgcta ctggtgctct tgaaggacag atgttccgga
 721   aaactgggag gcttatctca ctgagtgagc agaatctggt agactgctct gggcctcaag
 781   gcaatgaagg ctgcaatggt ggcctaatgg attatgcttt ccagtatgtt caggataatg
 841   gaggcctgga ctctgaggaa tcctatccat atgaggcaac agaagaatcc tgtaagtaca
 901   atcccaagta ttctgttgct aatgacaccg gctttgtgga catccctaag caggagaagg
 961   ccctgatgaa ggcagttgca actgtggggc ccatttctgt tgctattgat gcaggtcatg
1021   agtccttcct gttctataaa gaaggcattt attttgagcc agactgtagc agtgaagaca
1081   tggatcatgg tgtgctggtg gttggctacg gatttgaaag cacagaatca gataacaata
1141   aatattggct ggtgaagaac agctggggtg aagaatgggg catgggtggc tacgtaaaga
1201   tggccaaaga ccggagaaac cattgtggaa ttgcctcagc agccagctac cccactgtgt
1261   gagctggtgg acggtgatga ggaaggactt gactggggat ggcgcatgca tgggaggaat
1321   tcatcttcag tctaccagcc cccgctgtgt cggatacaca ctcgaatcat tgaagatccg
```

-continued

SEQUENCE LISTING

```
1381   agtgtgattt gaattctgtg atattttcac actggtaaat gttacctcta ttttaattac
1441   tgctataaat aggtttatat tattgattca cttactgact ttgcattttc gtttttaaaa
1501   ggatgtataa atttttacct gtttaaataa aatttaattt caaatgtagt ggtggggctt
1561   ctttctattt ttgatgcact gaatttttgt gtaataaaga acataattgg gctctaagcc
1621   ataaaa

SEQ ID NO: 61 Human Cathepsin L Polypeptide, variant 3
MNPTLILAAFCLGIASATLTFDHSLEAQWTKWKAMHNRLYGMNE
EGWRRAVWEKNMKMIELHNQEYREGKHSFTMAMNAFGDMTSEEFRQVMNGFQNRKPRK
GKVFQEPLFYEAPRSVDWREKGYVTPVKNQGQCGSCWAFSATGALEGQMFRKTGRLIS
LSEQNLVDCSGPQGNEGCNGGLMDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYS
VANDTGFVDIPKQEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDH
GVLVVGYGFESTESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV

SEQ ID NO: 62 Human Cathepsin L mRNA, variant 4
   1   ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc
  61   agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac
 121   cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag
 181   gacgcggtcg agttttaaaa catgaatcct acactcatcc ttgctgcctt ttgcctggga
 241   attgcctcag ctactctaac atttgatcac agtttagagg cacagtggac caagtggaag
 301   gcgatgcaca acagattata cggcatgaat gaagaaggat ggaggagagc agtgtgggag
 361   aagaacatga agatgattga actgcacaat caggaataca gggaagggaa acacagcttc
 421   acaatggcca tgaacgcctt tggagacatg accagtgaag aattcaggca ggtgatgaat
 481   ggctttcaaa accgtaagcc caggaagggg aaagtgttcc aggaacctct gttttatgag
 541   gcccccagat ctgtggattg gagagagaaa ggctacgtga ctcctgtgaa gaatcagggt
 601   cagtgtggtt cttgttgggc ttttagtgct actggtgctc ttgaaggaca gatgttccgg
 661   aaaactggga ggcttatctc actgagtgag cagaatctgg tagactgctc tgggcctcaa
 721   ggcaatgaag gctgcaatgg tggcctaatg gattatgctt tccagtatgt tcaggataat
 781   ggaggcctgg actctgagga atcctatcca tatgaggcaa cagaagaatc ctgtaagtac
 841   aatcccaagt attctgttgc taatgacacc ggctttgtgg acatccctaa gcaggagaag
 901   gccctgatga aggcagttgc aactgtgggg cccatttctg ttgctattga tgcaggtcat
 961   gagtccttcc tgttctataa agaaggcatt tattttgagc cagactgtag cagtgaagac
1021   atggatcatg gtgtgctggt ggttggctac ggatttgaaa gcacagaatc agataacaat
1081   aaatattggc tggtgaagaa cagctggggt gaagaatggg gcatgggtgg ctacgtaaag
1141   atggccaaag accggagaaa ccattgtgga attgcctcag cagccagcta ccccactgtg
1201   tgagctggtg gacggtgatg aggaaggact tgactgggga tggcgcatgc atgggaggaa
1261   ttcatcttca gtctaccagc ccccgctgtg tcggatacac actcgaatca ttgaagatcc
1321   gagtgtgatt tgaattctgt gatattttca cactggtaaa tgttacctct attttaatta
1381   ctgctataaa taggtttata ttattgattc acttactgac tttgcatttt cgtttttaaa
1441   aggatgtata aattttacc tgtttaaata aaatttaatt tcaaatgtag tggtggggct
1501   tctttctatt tttgatgcac tgaatttttg tgtaataaag aacataattg ggctctaagc
1561   cataaaa

SEQ ID NO: 63 Human Cathepsin L Polypeptide, variant 4
MNPTLILAAFCLGIASATLTFDHSLEAQWTKWKAMHNRLYGMNE
EGWRRAVWEKNMKMIELHNQEYREGKHSFTMAMNAFGDMTSEEFRQVMNGFQNRKPRK
GKVFQEPLFYEAPRSVDWREKGYVTPVKNQGQCGSCWAFSATGALEGQMFRKTGRLIS
LSEQNLVDCSGPQGNEGCNGGLMDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYS
VANDTGFVDIPKQEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDH
GVLVVGYGFESTESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV

SEQ ID NO: 64 Human Cathepsin L mRNA, variant 5
   1   ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc
  61   agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac
 121   cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag
 181   gacgcggtcg agtaggtttt aaaacatgaa tcctacactc atccttgctg ccttttgcct
 241   gggaattgcc tcagctactc taacatttga tcacagttta gaggcacagt ggaccaagtg
 301   gaaggctgca atggtggcct aatggattat gctttccagt atgttcagga taatggaggc
 361   ctggactctg aggaatccta tccatatgag gcaacagaag aatcctgtaa gtacaatccc
 421   aagtattctg ttgctaatga caccggcttt gtggacatcc ctaagcagga gaaggccctg
 481   atgaaggcag ttgcaactgt ggggcccatt tctgttgcta ttgatgcagg tcatgagtcc
 541   ttcctgttct ataaagaagg catttatttt gagccagact gtagcagtga agacatggat
 601   catggtgtgc tggtggttgg ctacggattt gaaagcacag aatcagataa caataaatat
 661   tggctggtga agaacagctg gggtgaagaa tggggcatgg gtggctacgt aaagatggcc
 721   aaagaccgga gaaaccattg tggaattgcc tcagcagcca gctaccccac tgtgtgagct
 781   ggtggacggt gatgaggaag gacttgactg gggatggcgc atgcatggga ggaattcatc
 841   ttcagtctac cagcccccgc tgtgtcggat acacactcga atcattgaag atccgagtgt
 901   gatttgaatt ctgtgatatt ttcacactgg taaatgttac ctctatttta attactgcta
 961   taaataggtt tatattattg attcacttac tgactttgca ttttcgtttt taaaaggatg
1021   tataaatttt tacctgttta aataaaattt aatttcaaat gtagtggtgg ggcttctttc
1081   tatttttgat gcactgaatt tttgtgtaat aaagaacata attgggctct aagccataaa
1141   a
```

-continued

SEQUENCE LISTING

SEQ ID NO: 65 Human Cathepsin L Polypeptide, variant 5
MDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYSVANDTGFV
DIPKQEKALMKAVATVGPISVAIDAGHESFLFYKEGIYFEPDCSSEDMDHGVLVVGYG
FESTESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV

SEQ ID NO: 66 Human Cathepsin L mRNA, variant 6
```
   1   acagctctgg acaggctgct tttcattttg gtgagtccat ccagtacctc cacgtgccct
  61   gtttttctcc aggcacatcc ttggcctctt ccacagtcct tgggttttaa aacatgaatc
 121   ctacactcat ccttgctgcc ttttgcctgg gaattgcctc agctactcta acatttgatc
 181   acagtttaga ggcacagtgg accaagtgga aggcgatgca caacagatta tacggcatga
 241   atgaagaagg atggaggaga gcagtgtggg agaagaacat gaagatgatt gaactgcaca
 301   atcaggaata cagggaaggg aaacacagct tcacaatggc catgaacgcc tttggagaca
 361   tgaccagtga agaattcagg caggtgatga atggctttca aaaccgtaag cccaggaagg
 421   ggaaagtgtt ccaggaacct ctgttttatg aggcccccag atctgtggat tggagagaga
 481   aaggctacgt gactcctgtg aagaatcagg gtcagtgtgg ttcttgttgg gcttttagtg
 541   ctactggtgc tcttgaagga cagatgttcc ggaaaactgg gaggcttatc tcactgagtg
 601   agcagaatct ggtagactgc tctgggcctc aaggcaatga aggctgcaat ggtggcctaa
 661   tggattatgc tttccagtat gttcaggata atggaggcct ggactctgag gaatcctatc
 721   catatgaggc aacagaagaa tcctgtaagt acaatcccaa gtattctgtt gctaatgaca
 781   ccggctttgt ggacatccct aagcaggaga aggccctgat gaaggcagtt gcaactgtgg
 841   ggcccatttc tgttgctatt gatgcaggtc atgagtcctt cctgttctat aaagaaggca
 901   tttattttga gccagactgt agcagtgaag acatggatca tggtgtgctg gtggttggct
 961   acggatttga aagcacagaa tcagataaca ataaatattg gctggtgaag aacagctggg
1021   gtgaagaatg gggcatgggt ggctacgtaa agatggccaa agaccggaga aaccattgtg
1081   gaattgcctc agcagccagc taccccactg tgtgagctgg tggacggtga tgaggaagga
1141   cttgactggg gatggcgcat gcatgggagg aattcatctt cagtctacca gcccccgctg
1201   tgtcggatac acactcgaat cattgaagat ccgagtgtga tttgaattct gtgatatttt
1261   cacactggta aatgttacct ctattttaat tactgctata aataggttta tattattgat
1321   tcacttactg actttgcatt ttcgttttta aaaggatgta taaatttttа cctgtttaaa
1381   taaaatttaa tttcaaatgt a
```

SEQ ID NO: 67 Human Cathepsin L Polypeptide, variant 6
MNPTLILAAFCLGIASATLTFDHSLEAQWTKWKAMHNRLYGMNE
EGWRRAVWEKNMKMIELHNQEYREGKHSFTMAMNAFGDMTSEEFRQVMNGFQNRKPRK
GKVFQEPLFYEAPRSVDWREKGYVTPVKNQGQCGSCWAFSATGALEGQMFRKTGRLIS
LSEQNLVDCSGPQGNEGCNGGLMDYAFQYVQDNGGLDSEESYPYEATEESCKYNPKYS
VANDTGFVDIPKQEKALMKAVATVGPISVAIDAGHESELFYKEGIYFEPDCSSEDMDH
GVLVVGYGFESTESDNNKYWLVKNSWGEEWGMGGYVKMAKDRRNHCGIASAASYPTV

SEQ ID NO: 68 Human Cathepsin D Polypeptide
MQPSSLLPLALCLLAAPASALVRIPLHKFTSIRRTMSEVGGSVEDLIAKGPVSKYSQAVP
AVTEGPIPEVLKNYMDAQYYGEIGIGTPPQCFTVVFDTGSSNLWVPSIHCKLLDIACWIH
HKYNSDKSSTYVKNGTSFDIHYGSGSLSGYLSQDTVSVPCQSASSASALGGVKVERQVFG
EATKQPGITFIAAKEDGILGMAYPRISVNNVLPVEDNLMQQKLVDQNIFSFYLSRDPDAQ
PGGELMLGGTDSKYYKGSLSYLNVTRKAYWQVHLDQVEVASGLTLCKEGCEAIVDTGTSL
MVGPVDEVRELQKAIGAVPLIQGEYMIPCEKVSTLPAITLKLGGKGYKLSPEDYTLKVSQ
AGKTLCLSGFMGMDIPPPSGPLWILGDVFIGRYYTVFDRDNNRVGFAEAARL

SEQ ID NO: 69 Human Cathepsin E Polypeptide, Isoform 3
MKTLLLLLLVLLELGEAQGSLHRVPLRRHPSLKKKLRARSQLSEFWKSHNLDMIQFTESC
SMDQSAKEPLINYLDMEYFGTISIGSPPQNFTVIFDTGSSNLWVPSVYCTSPACKTHSRF
QPSQSSTYSQPGQSFSIQYGTGSLSGIIGADQVSAFATQVEGLTVVGQQFGESVTEPGQT
FVDAEFDGILGLGYPSLAVGGVTPVFDNMMAQNLVDLPMFSVYMSSNPEGGAGSELIFGG
YDHSHFSGSLNWVPVTKQAYWQIALDNIQVGGTVMFCSEGCQAIVDTGTSLITGPSDKIK
QLQNAIGAAPVDGEYAVECANLNVMPDVTFTINGVPYTLSPTAYTLLDFVDGMQFCSSGF
QGLDIHPPAGPLWILGDVFIRQFYSVFDRGNNRVGLAPAVP

SEQ ID NO: 70 Human Cathepsin E Polypeptide, Isoform 1
MKTLLLLLLVLLELGEAQGSLHRVPLRRHPSLKKKLRARSQLSEFWKSHNLDMIQFTESC
SMDQSAKEPLINYLDMEYFGTISIGSPPQNFTVIFDTGSSNLWVPSVYCTSPACKTHSRF
QPSQSSTYSQPGQSFSIQYGTGSLSGIIGADQVSVEGLTVVGQQFGESVTEPGQTFVDAE
FDGILGLGYPSLAVGGVTPVFDNMMAQNLVDLPMFSVYMSSNPEGGAGSELIFGGYDHSH
FSGSLNWVPVTKQAYWQIALDNIQVGGTVMFCSEGCQAIVDTGTSLITGPSDKIKQLQNA
IGAAPVDGEYAVECANLNVMPDVTFTINGVPYTLSPTAYTLLDFVDGMQFCSSGFQGLDI
HPPAGPLWILGDVFIRQFYSVFDRGNNRVGLAPAVP

SEQ ID NO: 71 Human Cathepsin E Polypeptide, Isoform 2
MKTLLLLLLVLLELGEAQGSLHRVPLRRHPSLKKKLRARSQLSEFWKSHNLDMIQFTESC
SMDQSAKEPLINYLDMEYFGTISIGSPPQNFTVIFDTGSSNLWVPSVYCTSPACKTHSRF
QPSQSSTYSQPGQSFSIQYGTGSLSGIIGADQVSVEGLTVVGQQFGESVTEPGQTFVDAE
FDGILGLGYPSLAVGGVTPVFDNMMAQNLVDLPMFSVYMSSNPEGGAGSELIFGGYDHSH
FSGSLNWVPVTKQAYWQIALDNMLWSVPTLTSCRMSPSPLTESPIPSAQLPTPYWTSWME
CSSAAVAFKDLTSTLQLGPSGSWGMSSFDSFTQSLTVGITVWDWPQQSPKEGPCVCACLS
DRP

-continued

SEQUENCE LISTING

SEQ ID NO: 72 cell permeable peptide, L803-mts
GKEAPPAPPQSP

SEQUENCE LISTING

<160> NUMBER OF SEQ ID NOS: 72

<210> SEQ ID NO 1
<211> LENGTH: 2254
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 1

agagtgcacc cgaatccacg ggctcggagg cagcagccat ctctcggcca tagggcaggc 60

cagctggcgc cgggggctat tttgggcggc gggcaatgat ggtgaccgca aggcgacctt 120

gtaaggcatt tccccccctga ctcccttccc cgagcctctg cccgggggtc ctagcgccgc 180

tttctcagcc atcccgccta caacttagcc gtccacaaca ggatcatctg atcgcgtgcg 240

cccgggctac gatctgcgag gcccgcggac cttgacccgg cattgaccgc caccgccccc 300

caggtccgta gggaccaaag aaggggcggg aggaagactg tcacgtggcg ccggagttca 360

cgtgactcgt acacatgact tccagtcccc gggcgcctcc tggagagcaa ggacgcgggg 420

gagcagagat gatccgagcc gcgccgccgc cgctgttcct gctgctgctg ctgctgctgc 480

tgctagtgtc ctgggcgtcc cgaggcgagg cagcccccga ccaggacgag atccagcgcc 540

tccccgggct ggccaagcag ccgtctttcc gccagtactc cggctacctc aaaggctccg 600

gctccaagca cctccactac tggtttgtgg agtcccagaa ggatcccgag aacagccctg 660

tggtgctttg gctcaatggg ggtcccggct gcagctcact agatgggctc ctcacagagc 720

atggcccctt cctggtccag ccagatggtg tcaccctgga gtacaacccc tattcttgga 780

atctgattgc caatgtgtta tacctggagt ccccagctgg ggtgggcttc tcctactccg 840

atgacaagtt ttatgcaact aatgacactg aggtcgccca gagcaatttt gaggcccttc 900

aagatttctt ccgcctcttt ccggagtaca agaacaacaa acttttcctg accggggaga 960

gctatgctgg catctacatc cccaccctgg ccgtgctggt catgcaggat cccagcatga 1020

accttcaggg gctggctgtg ggcaatggac tctcctccta tgagcagaat gacaactccc 1080

tggtctactt tgcctactac catggccttc tggggaacag gctttggtct tctctccaga 1140

cccactgctg ctctcaaaac aagtgtaact tctatgacaa caaagacctg gaatgcgtga 1200

ccaatcttca ggaagtggcc cgcatcgtgg gcaactctgg cctcaacatc tacaatctct 1260

atgccccgtg tgctggaggg gtgcccagcc attttaggta tgagaaggac actgttgtgg 1320

tccaggattt gggcaacatc ttcactcgcc tgccactcaa gcggatgtgg catcaggcac 1380

tgctgcgctc aggggataaa gtgcgcatgg acccccccctg caccaacaca acagctgctt 1440

ccacctacct caacaacccg tacgtgcgga aggccctcaa catcccggag cagctgccac 1500

aatgggacat gtgcaacttt ctggtaaact tacagtaccg ccgtctctac cgaagcatga 1560

actcccagta tctgaagctg cttagctcac agaaatacca gatcctatta tataatggag 1620

atgtagacat ggcctgcaat ttcatggggg atgagtggtt tgtggattcc ctcaaccaga 1680

-continued

agatggaggt gcagcgccgg ccctggttag tgaagtacgg ggacagcggg gagcagattg 1740

ccggcttcgt gaaggagttc tcccacatcg cctttctcac gatcaagggc gccggccaca 1800

tggttcccac cgacaagccc ctcgctgcct tcaccatgtt ctcccgcttc ctgaacaagc 1860

agccatactg atgaccacag caaccagctc cacggcctga tgcagcccct cccagcctct 1920

cccgctagga gagtcctctt ctaagcaaag tgcccctgca ggccgggttc tgccgccagg 1980

actgcccccc tcccagagcc ctgtacatcc cagactgggc ccagggtctc ccatagacag 2040

cctgggggca agttagcact ttattcccgc agcagttcct gaatggggtg gcctggcccc 2100

ttctctgctt aaagaatgcc ctttatgatg cactgattcc atcccaggaa cccaacagag 2160

ctcaggacag cccacaggga ggtggtggac ggactgtaat tgatagattg attatggaat 2220

taaattgggt acagcttcaa aaaaaaaaaa aaaa 2254

<210> SEQ ID NO 2
<211> LENGTH: 498
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 2

Met Thr Ser Ser Pro Arg Ala Pro Pro Gly Glu Gln Gly Arg Gly Gly
1 5 10 15

Ala Glu Met Ile Arg Ala Ala Pro Pro Pro Leu Phe Leu Leu Leu Leu
20 25 30

Leu Leu Leu Leu Leu Val Ser Trp Ala Ser Arg Gly Glu Ala Ala Pro
35 40 45

Asp Gln Asp Glu Ile Gln Arg Leu Pro Gly Leu Ala Lys Gln Pro Ser
50 55 60

Phe Arg Gln Tyr Ser Gly Tyr Leu Lys Gly Ser Gly Ser Lys His Leu
65 70 75 80

His Tyr Trp Phe Val Glu Ser Gln Lys Asp Pro Glu Asn Ser Pro Val
85 90 95

Val Leu Trp Leu Asn Gly Gly Pro Gly Cys Ser Ser Leu Asp Gly Leu
100 105 110

Leu Thr Glu His Gly Pro Phe Leu Val Gln Pro Asp Gly Val Thr Leu
115 120 125

Glu Tyr Asn Pro Tyr Ser Trp Asn Leu Ile Ala Asn Val Leu Tyr Leu
130 135 140

Glu Ser Pro Ala Gly Val Gly Phe Ser Tyr Ser Asp Asp Lys Phe Tyr
145 150 155 160

Ala Thr Asn Asp Thr Glu Val Ala Gln Ser Asn Phe Glu Ala Leu Gln
165 170 175

Asp Phe Phe Arg Leu Phe Pro Glu Tyr Lys Asn Asn Lys Leu Phe Leu
180 185 190

Thr Gly Glu Ser Tyr Ala Gly Ile Tyr Ile Pro Thr Leu Ala Val Leu
195 200 205

Val Met Gln Asp Pro Ser Met Asn Leu Gln Gly Leu Ala Val Gly Asn
210 215 220

Gly Leu Ser Ser Tyr Glu Gln Asn Asp Asn Ser Leu Val Tyr Phe Ala
225 230 235 240

Tyr Tyr His Gly Leu Leu Gly Asn Arg Leu Trp Ser Ser Leu Gln Thr
245 250 255

-continued

```
His Cys Cys Ser Gln Asn Lys Cys Asn Phe Tyr Asp Asn Lys Asp Leu
            260                 265                 270

Glu Cys Val Thr Asn Leu Gln Glu Val Ala Arg Ile Val Gly Asn Ser
        275                 280                 285

Gly Leu Asn Ile Tyr Asn Leu Tyr Ala Pro Cys Ala Gly Gly Val Pro
    290                 295                 300

Ser His Phe Arg Tyr Glu Lys Asp Thr Val Val Val Gln Asp Leu Gly
305                 310                 315                 320

Asn Ile Phe Thr Arg Leu Pro Leu Lys Arg Met Trp His Gln Ala Leu
                325                 330                 335

Leu Arg Ser Gly Asp Lys Val Arg Met Asp Pro Pro Cys Thr Asn Thr
            340                 345                 350

Thr Ala Ala Ser Thr Tyr Leu Asn Asn Pro Tyr Val Arg Lys Ala Leu
        355                 360                 365

Asn Ile Pro Glu Gln Leu Pro Gln Trp Asp Met Cys Asn Phe Leu Val
    370                 375                 380

Asn Leu Gln Tyr Arg Arg Leu Tyr Arg Ser Met Asn Ser Gln Tyr Leu
385                 390                 395                 400

Lys Leu Leu Ser Ser Gln Lys Tyr Gln Ile Leu Leu Tyr Asn Gly Asp
                405                 410                 415

Val Asp Met Ala Cys Asn Phe Met Gly Asp Glu Trp Phe Val Asp Ser
            420                 425                 430

Leu Asn Gln Lys Met Glu Val Gln Arg Arg Pro Trp Leu Val Lys Tyr
        435                 440                 445

Gly Asp Ser Gly Glu Gln Ile Ala Gly Phe Val Lys Glu Phe Ser His
    450                 455                 460

Ile Ala Phe Leu Thr Ile Lys Gly Ala Gly His Met Val Pro Thr Asp
465                 470                 475                 480

Lys Pro Leu Ala Ala Phe Thr Met Phe Ser Arg Phe Leu Asn Lys Gln
                485                 490                 495

Pro Tyr
```

<210> SEQ ID NO 3
<211> LENGTH: 2088
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 3

```
gagctacttg aagaccaatt agagtccggg aagcgcggcg gggcctccag accggggcgg     60

gcttaagggt gacatctgcg ctttaaaggg tccgggtcag ctgactcccg actctgtgga    120

gtctagctgc cagggtcgcg gcagctgcgg ggagagatga ctggggagcg acccagcacg    180

gcgctcccgg acagacgctg ggggccgcgg attctgggct tctggggagg ctgtagggtt    240

tgggtgtttg ccgcgatctt cctgctgctg tctctggcag cctcctggtc caaggctgag    300

aacgacttcg gtctggtgca gccgctggtg accatggagc aactgctgtg ggtgagcggg    360

agacagatcg gctcagtgga caccttccgc atcccgctca tcacagccac tccgcggggc    420

actcttctcg cctttgctga ggcgaggaaa atgtcctcat ccgatgaggg ggccaagttc    480

atcgccctgc ggaggtccat ggaccagggc agcacatggt ctcctacagc gttcattgtc    540

aatgatgggg atgtccccga tgggctgaac cttggggcag tagtgagcga tgttgagaca    600

ggagtagtat ttcttttcta ctccctttgt gctcacaagg ccggctgcca ggtggcctct    660
```

-continued

accatgttgg tatggagcaa ggatgatggt gtttcctgga gcacaccccg gaatctctcc 720

ctggatattg gcactgaagt gtttgcccct ggaccgggct ctggtattca gaaacagcgg 780

gagccacgga agggccgcct catcgtgtgt ggccatggga cgctggagcg ggacggagtc 840

ttctgtctcc tcagcgatga tcatggtgcc tcctggcgct acggaagtgg ggtcagcggc 900

atcccctacg gtcagcccaa gcaggaaaat gatttcaatc ctgatgaatg ccagccctat 960

gagctcccag atggctcagt cgtcatcaat gcccgaaacc agaacaacta ccactgccac 1020

tgccgaattg tcctccgcag ctatgatgcc tgtgatacac taaggccccg tgatgtgacc 1080

ttcgaccctg agctcgtgga ccctgtggta gctgcaggag ctgtagtcac cagctccggc 1140

attgtcttct tctccaaccc agcacatcca gagttccgag tgaacctgac cctgcgatgg 1200

agcttcagca atggtacctc atggcggaaa gagacagtcc agctatggcc aggccccagt 1260

ggctattcat ccctggcaac cctggagggc agcatggatg gagaggagca ggccccccag 1320

ctctacgtcc tgtatgagaa aggccggaac cactacacag agagcatctc cgtggccaaa 1380

atcagtgtct atgggacact ctgagctgtg ccactgccac aggggtattc tgccttcagg 1440

actctgcctt caggaacacg ggtctgtaga gggtctgctg gagacgcctg aaagacagtt 1500

ccatcttcct ttagactcca gccttggcaa aatcaccttc cctttaccag ggaaatcact 1560

tcctttagga ctgaaagcta ggcgtcctct cccacaaaaa agtcctgccc tcatctgaga 1620

atactgtctt tccatatggc taagtgtggc cccaccaccc tctctgccct cccgggacat 1680

tgattggtcc tgtcttgggc aggtctagtg agctgtagaa ttgaatcaat gtgaactcag 1740

ggaactgggg aaggctgagc ctcctctttg gtgttgcggt aagataaccg acagggctgg 1800

tgaaagtccc cagatggcag gatatttggt ttcagagtaa ggactaggtg caccaccatg 1860

actgactatc aatcaaaatg tttgtaactt aaaatttta atgaaggata atgaatattt 1920

gtagagtctc tatggttctg tcaatgcaca tcttcgtgtc tgttttcctc atgtatcctt 1980

gtgagcctgg gtgagttctg gggagagacc tgatgtgcgt actgcctgtg aaaatctgac 2040

tttggcaaat caaatcctct tttccttttg aaaaaaaaaa aaaaaaaa 2088

<210> SEQ ID NO 4
<211> LENGTH: 415
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 4

Met Thr Gly Glu Arg Pro Ser Thr Ala Leu Pro Asp Arg Arg Trp Gly
1 5 10 15

Pro Arg Ile Leu Gly Phe Trp Gly Gly Cys Arg Val Trp Val Phe Ala
20 25 30

Ala Ile Phe Leu Leu Leu Ser Leu Ala Ala Ser Trp Ser Lys Ala Glu
35 40 45

Asn Asp Phe Gly Leu Val Gln Pro Leu Val Thr Met Glu Gln Leu Leu
50 55 60

Trp Val Ser Gly Arg Gln Ile Gly Ser Val Asp Thr Phe Arg Ile Pro
65 70 75 80

Leu Ile Thr Ala Thr Pro Arg Gly Thr Leu Leu Ala Phe Ala Glu Ala
85 90 95

Arg Lys Met Ser Ser Ser Asp Glu Gly Ala Lys Phe Ile Ala Leu Arg

-continued

```
            100                 105                 110

Arg Ser Met Asp Gln Gly Ser Thr Trp Ser Pro Thr Ala Phe Ile Val
        115                 120                 125

Asn Asp Gly Asp Val Pro Asp Gly Leu Asn Leu Gly Ala Val Val Ser
    130                 135                 140

Asp Val Glu Thr Gly Val Val Phe Leu Phe Tyr Ser Leu Cys Ala His
145                 150                 155                 160

Lys Ala Gly Cys Gln Val Ala Ser Thr Met Leu Val Trp Ser Lys Asp
                165                 170                 175

Asp Gly Val Ser Trp Ser Thr Pro Arg Asn Leu Ser Leu Asp Ile Gly
            180                 185                 190

Thr Glu Val Phe Ala Pro Gly Pro Gly Ser Gly Ile Gln Lys Gln Arg
        195                 200                 205

Glu Pro Arg Lys Gly Arg Leu Ile Val Cys Gly His Gly Thr Leu Glu
    210                 215                 220

Arg Asp Gly Val Phe Cys Leu Leu Ser Asp Asp His Gly Ala Ser Trp
225                 230                 235                 240

Arg Tyr Gly Ser Gly Val Ser Gly Ile Pro Tyr Gly Gln Pro Lys Gln
                245                 250                 255

Glu Asn Asp Phe Asn Pro Asp Glu Cys Gln Pro Tyr Glu Leu Pro Asp
            260                 265                 270

Gly Ser Val Val Ile Asn Ala Arg Asn Gln Asn Asn Tyr His Cys His
        275                 280                 285

Cys Arg Ile Val Leu Arg Ser Tyr Asp Ala Cys Asp Thr Leu Arg Pro
    290                 295                 300

Arg Asp Val Thr Phe Asp Pro Glu Leu Val Asp Pro Val Val Ala Ala
305                 310                 315                 320

Gly Ala Val Val Thr Ser Ser Gly Ile Val Phe Phe Ser Asn Pro Ala
                325                 330                 335

His Pro Glu Phe Arg Val Asn Leu Thr Leu Arg Trp Ser Phe Ser Asn
            340                 345                 350

Gly Thr Ser Trp Arg Lys Glu Thr Val Gln Leu Trp Pro Gly Pro Ser
        355                 360                 365

Gly Tyr Ser Ser Leu Ala Thr Leu Glu Gly Ser Met Asp Gly Glu Glu
    370                 375                 380

Gln Ala Pro Gln Leu Tyr Val Leu Tyr Glu Lys Gly Arg Asn His Tyr
385                 390                 395                 400

Thr Glu Ser Ile Ser Val Ala Lys Ile Ser Val Tyr Gly Thr Leu
                405                 410                 415


<210> SEQ ID NO 5
<211> LENGTH: 3540
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 5

ggtggtggaa tatagagctc atgtgatccg tcacatgaca gcagatccgc ggaagggcag     60

aatgggactc caagcctgcc tcctagggct ctttgccctc atcctctctg gcaaatgcag    120

ttacagcccg gagcccgacc agcggaggac gctgccccca ggctgggtgt ccctgggccg    180

tgcggaccct gaggaagagc tgagtctcac ctttgccctg agacagcaga atgtggaaag    240

actctcggag ctggtgcagg ctgtgtcgga tcccagctct cctcaatacg gaaaatacct    300
```

-continued

gaccctagag aatgtggctg atctggtgag gccatcccca ctgaccctcc acacggtgca 360

aaaatggctc ttggcagccg gagcccagaa gtgccattct gtgatcacac aggactttct 420

gacttgctgg ctgagcatcc gacaagcaga gctgctgctc cctggggctg agtttcatca 480

ctatgtggga ggacctacgg aaacccatgt tgtaaggtcc ccacatccct accagcttcc 540

acaggccttg gccccccatg tggactttgt ggggggactg caccgttttc ccccaacatc 600

atccctgagg caacgtcctg agccgcaggt gacagggact gtaggcctgc atctgggggt 660

aacccccctct gtgatccgta agcgatacaa cttgacctca caagacgtgg gctctggcac 720

cagcaataac agccaagcct gtgcccagtt cctggagcag tatttccatg actcagacct 780

ggctcagttc atgcgcctct tcggtggcaa ctttgcacat caggcatcag tagcccgtgt 840

ggttggacaa cagggccggg gccgggccgg gattgaggcc agtctagatg tgcagtacct 900

gatgagtgct ggtgccaaca tctccacctg ggtctacagt agccctggcc ggcatgaggg 960

acaggagccc ttcctgcagt ggctcatgct gctcagtaat gagtcagccc tgccacatgt 1020

gcatactgtg agctatggag atgatgagga ctccctcagc agcgcctaca tccagcgggt 1080

caacactgag ctcatgaagg ctgccgctcg gggtctcacc ctgctcttcg cctcaggtga 1140

cagtggggcc gggtgttggt ctgtctctgg aagacaccag ttccgcccta ccttccctgc 1200

ctccagcccc tatgtcacca cagtgggagg cacatccttc caggaacctt tcctcatcac 1260

aaatgaaatt gttgactata tcagtggtgg tggcttcagc aatgtgttcc cacggccttc 1320

ataccaggag gaagctgtaa cgaagttcct gagctctagc ccccacctgc caccatccag 1380

ttacttcaat gccagtggcc gtgcctaccc agatgtggct gcactttctg atggctactg 1440

ggtggtcagc aacagagtgc ccattccatg ggtgtccgga acctcggcct ctactccagt 1500

gtttgggggg atcctatcct tgatcaatga gcacaggatc cttagtggcc gccccctct 1560

tggctttctc aacccaaggc tctaccagca gcatggggca ggactctttg atgtaacccg 1620

tggctgccat gagtcctgtc tggatgaaga ggtagagggc cagggtttct gctctggtcc 1680

tggctgggat cctgtaacag gctggggaac acccaacttc ccagctttgc tgaagactct 1740

actcaacccc tgacccttc ctatcaggag agatggcttg tcccctgccc tgaagctggc 1800

agttcagtcc cttattctgc cctgttggaa gccctgctga accctcaact attgactgct 1860

gcagacagct tatctcccta accctgaaat gctgtgagct tgacttgact cccaacccta 1920

ccatgctcca tcatactcag gtctccctac tcctgcctta gattcctcaa taagatgctg 1980

taactagcat tttttgaatg cctctccctc cgcatctcat ctttctcttt tcaatcaggc 2040

ttttccaaag ggttgtatac agactctgtg cactatttca cttgatattc attccccaat 2100

tcactgcaag gagacctcta ctgtcaccgt ttactctttc ctaccctgac atccagaaac 2160

aatggcctcc agtgcatact tctcaatctt tgctttatgg cctttccatc atagttgccc 2220

actccctctc cttacttagc ttccaggtct taacttctct gactactctt gtcttcctct 2280

ctcatcaatt tctgcttctt catggaatgc tgaccttcat tgctccattt gtagattttt 2340

gctcttctca gtttactcat tgtcccctgg aacaaatcac tgacatctac aaccattacc 2400

atctcactaa ataagacttt ctatccaata atgattgata cctcaaatgt aagatgcgtg 2460

atactcaaca tttcatcgtc caccttccca accccaaaca attccatctc gtttcttctt 2520

ggtaaatgat gctatgcttt ttccaaccaa gccagaaacc tgtgtcatct tttcacccca 2580

-continued

ccttcaatca acaagtcctc aatcaacaag tcctactgac tgcacatctt aaatatatct 2640

ttatcagtcc acaagtcctt ccaattatat ttcccaagta tatctagaac ttatccactt 2700

atatccccac tgctactacc ttagtttagg gctatattct cttgaaaaaa agtgtcctta 2760

cttcctgcca atccccaagt catcttccag agtaaaatgc aaatcccatc aggccacttg 2820

gatgaaaacc cttcaaggat tactggatag aattcaggct ttcccctcca gcccccaatc 2880

atagctcaca aaccttcctt gctatttgtt cttaagtaaa aaatcatttt tcctcctccc 2940

tccccaaacc ccaaggaact ctcactcttg ctcaagctgt tccgtcccct taccaccccc 3000

gatacaactg ccaggttaat ttccagaatt cttgcaagac tcagttcaga agtcaccttc 3060

tttcgtgaat gttttgattc cctgaggcta ctttattttg gtatggctga aaaatcctag 3120

attttctaaa caaaacctgt ttgaatcttg gttctgatat ggactaggag agagactggg 3180

tcaagtaagc ttatctccct gaggctgttt cctcgtctgt taagtgtgaa tatcaatacc 3240

tgcctttcat aatcaccagg gaataaagtg gaataatgtt gataacagtg cttggcacct 3300

ggaagtaggt ggcagatgtt aacgcccttc ctcccttgca ctgcgccccc tgtgcctacc 3360

tctagcattg taacgaccac gtagtattga aatggccagt ttacttgtct gccttccttt 3420

ccaagaccgt tggtgcctag aggactagaa tcgtgtccta tttaactttg tgttcccagg 3480

tcctagctca ggagttggca aataagaatt aaatgtctgc tacaccgaaa accaaaaaaa 3540

<210> SEQ ID NO 6
<211> LENGTH: 563
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 6

Met Gly Leu Gln Ala Cys Leu Leu Gly Leu Phe Ala Leu Ile Leu Ser
1 5 10 15

Gly Lys Cys Ser Tyr Ser Pro Glu Pro Asp Gln Arg Arg Thr Leu Pro
20 25 30

Pro Gly Trp Val Ser Leu Gly Arg Ala Asp Pro Glu Glu Glu Leu Ser
35 40 45

Leu Thr Phe Ala Leu Arg Gln Gln Asn Val Glu Arg Leu Ser Glu Leu
50 55 60

Val Gln Ala Val Ser Asp Pro Ser Ser Pro Gln Tyr Gly Lys Tyr Leu
65 70 75 80

Thr Leu Glu Asn Val Ala Asp Leu Val Arg Pro Ser Pro Leu Thr Leu
85 90 95

His Thr Val Gln Lys Trp Leu Leu Ala Ala Gly Ala Gln Lys Cys His
100 105 110

Ser Val Ile Thr Gln Asp Phe Leu Thr Cys Trp Leu Ser Ile Arg Gln
115 120 125

Ala Glu Leu Leu Leu Pro Gly Ala Glu Phe His His Tyr Val Gly Gly
130 135 140

Pro Thr Glu Thr His Val Val Arg Ser Pro His Pro Tyr Gln Leu Pro
145 150 155 160

Gln Ala Leu Ala Pro His Val Asp Phe Val Gly Gly Leu His Arg Phe
165 170 175

Pro Pro Thr Ser Ser Leu Arg Gln Arg Pro Glu Pro Gln Val Thr Gly
180 185 190

-continued

```
Thr Val Gly Leu His Leu Gly Val Thr Pro Ser Val Ile Arg Lys Arg
        195                 200                 205

Tyr Asn Leu Thr Ser Gln Asp Val Gly Ser Gly Thr Ser Asn Asn Ser
    210                 215                 220

Gln Ala Cys Ala Gln Phe Leu Glu Gln Tyr Phe His Asp Ser Asp Leu
225                 230                 235                 240

Ala Gln Phe Met Arg Leu Phe Gly Gly Asn Phe Ala His Gln Ala Ser
                245                 250                 255

Val Ala Arg Val Val Gly Gln Gln Gly Arg Gly Arg Ala Gly Ile Glu
            260                 265                 270

Ala Ser Leu Asp Val Gln Tyr Leu Met Ser Ala Gly Ala Asn Ile Ser
        275                 280                 285

Thr Trp Val Tyr Ser Ser Pro Gly Arg His Glu Gly Gln Glu Pro Phe
    290                 295                 300

Leu Gln Trp Leu Met Leu Leu Ser Asn Glu Ser Ala Leu Pro His Val
305                 310                 315                 320

His Thr Val Ser Tyr Gly Asp Asp Glu Asp Ser Leu Ser Ser Ala Tyr
                325                 330                 335

Ile Gln Arg Val Asn Thr Glu Leu Met Lys Ala Ala Ala Arg Gly Leu
            340                 345                 350

Thr Leu Leu Phe Ala Ser Gly Asp Ser Gly Ala Gly Cys Trp Ser Val
        355                 360                 365

Ser Gly Arg His Gln Phe Arg Pro Thr Phe Pro Ala Ser Ser Pro Tyr
    370                 375                 380

Val Thr Thr Val Gly Gly Thr Ser Phe Gln Glu Pro Phe Leu Ile Thr
385                 390                 395                 400

Asn Glu Ile Val Asp Tyr Ile Ser Gly Gly Gly Phe Ser Asn Val Phe
                405                 410                 415

Pro Arg Pro Ser Tyr Gln Glu Glu Ala Val Thr Lys Phe Leu Ser Ser
            420                 425                 430

Ser Pro His Leu Pro Pro Ser Ser Tyr Phe Asn Ala Ser Gly Arg Ala
        435                 440                 445

Tyr Pro Asp Val Ala Ala Leu Ser Asp Gly Tyr Trp Val Val Ser Asn
    450                 455                 460

Arg Val Pro Ile Pro Trp Val Ser Gly Thr Ser Ala Ser Thr Pro Val
465                 470                 475                 480

Phe Gly Gly Ile Leu Ser Leu Ile Asn Glu His Arg Ile Leu Ser Gly
                485                 490                 495

Arg Pro Pro Leu Gly Phe Leu Asn Pro Arg Leu Tyr Gln Gln His Gly
            500                 505                 510

Ala Gly Leu Phe Asp Val Thr Arg Gly Cys His Glu Ser Cys Leu Asp
        515                 520                 525

Glu Glu Val Glu Gly Gln Gly Phe Cys Ser Gly Pro Gly Trp Asp Pro
    530                 535                 540

Val Thr Gly Trp Gly Thr Pro Asn Phe Pro Ala Leu Leu Lys Thr Leu
545                 550                 555                 560

Leu Asn Pro
```

<210> SEQ ID NO 7
<211> LENGTH: 3783
<212> TYPE: DNA

-continued

<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 7

ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg 60

ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct 120

ctcggctgca gcgctgggtg gatctaggat ccggcttcca acatgtggca gctctgggcc 180

tccctctgct gcctgctggt gttggccaat gcccggagca ggccctcttt ccatcccctg 240

tcggatgagc tggtcaacta tgtcaacaaa cggaatacca cgtggcaggc cgggcacaac 300

ttctacaacg tggacatgag ctacttgaag aggctatgtg gtaccttcct gggtgggccc 360

aagccacccc agagagttat gtttaccgag gacctgaagc tgcctgcaag cttcgatgca 420

cgggaacaat ggccacagtg tcccaccatc aaagagatca gagaccaggg ctcctgtggc 480

tcctgctggg ccttcggggc tgtggaagcc atctctgacc ggatctgcat ccacaccaat 540

gcgcacgtca gcgtggaggt gtcggcggag gacctgctca catgctgtgg cagcatgtgt 600

ggggacggct gtaatggtgg ctatcctgct gaagcttgga acttctggac aagaaaaggc 660

ctggtttctg gtggcctcta tgaatcccat gtagggtgca gaccgtactc catccctccc 720

tgtgagcacc acgtcaacgg ctcccggccc ccatgcacgg gggagggaga tacccccaag 780

tgtagcaaga tctgtgagcc tggctacagc ccgacctaca aacaggacaa gcactacgga 840

tacaattcct acagcgtctc caatagcgag aaggacatca tggccgagat ctacaaaaac 900

ggccccgtgg agggagcttt ctctgtgtat tcggacttcc tgctctacaa gtcaggagtg 960

taccaacacg tcaccggaga gatgatgggt ggccatgcca tccgcatcct gggctgggga 1020

gtggagaatg gcacacccta ctggctggtt gccaactcct ggaacactga ctggggtgac 1080

aatggcttct ttaaaatact cagaggacag gatcactgtg gaatcgaatc agaagtggtg 1140

gctggaattc cacgcaccga tcagtactgg gaaaagatct aatctgccgt gggcctgtcg 1200

tgccagtcct gggggcgaga tcggggtaga aatgcatttt attctttaag ttcacgtaag 1260

atacaagttt cagacagggt ctgaaggact ggattggcca aacatcagac ctgtcttcca 1320

aggagaccaa gtcctggcta catcccagcc tgtggttaca gtgcagacag gccatgtgag 1380

ccaccgctgc cagcacagag cgtccttccc cctgtagact agtgccgtag ggagtacctg 1440

ctgccccagc tgactgtggc cccctccgtg atccatccat ctccagggag caagacagag 1500

acgcaggaat ggaaagcgga gttcctaaca ggatgaaagt tcccccatca gttcccccag 1560

tacctccaag caagtagctt tccacatttg tcacagaaat cagaggagag acggtgttgg 1620

gagccctttg gagaacgcca gtctcccagg ccccctgcat ctatcgagtt tgcaatgtca 1680

caacctctct gatcttgtgc tcagcatgat tctttaatag aagttttatt ttttcgtgca 1740

ctctgctaat catgtgggtg agccagtgga acagcgggag acctgtgcta gttttacaga 1800

ttgcctcctt atgacgcggc tcaaaaggaa accaagtggt caggagttgt ttctgaccca 1860

ctgatctcta ctaccacaag gaaaatagtt taggagaaac cagcttttac tgtttttgaa 1920

aaattacagc ttcaccctgt caagttaaca aggaatgcct gtgccaataa aagttttctc 1980

caacttgaag tctactctga tgggatctca gatcctttgt cactgcctat agacttgtag 2040

ctgctgtctc tctttgtccc tgcagagaat cacgtcctgg aactgcatgt tcttgcgact 2100

cttgggactt catcttaact tctcgctgcc ccagccatgt tttcaaccat ggcatccctc 2160

-continued

ccccaattag ttccctgtca tcctcgtcaa ccttctctgt aagtgcctgg taagcttgcc 2220

cttgcttaag aactcaaaac atagctgtgc tctatttttt tgttgttgtt gtgactgaca 2280

gagtgagatt ccgtctccca ggctggagtg cagtggcgcc ttctcagctc actgcaacct 2340

gcagcctcct agattcaagc gattctcctg cttcagcctt ccgagtagct gggatgacag 2400

gcactcacca atatgcctgg gtaatttttg tatttttaag tacatacagg atttcaccat 2460

gttggccagg ctagtttcaa actcccggcc tcaggtggtc tgcctgcctc agcctcccaa 2520

agtgttggga ttacaggcgt gagccactgg gccctgcctg tatttttat cagccacaaa 2580

tccagcaaca agctgaggat tcagctcata aaacaggctt ggtgtcttgg tgatctcaca 2640

taaccaagat gctaccccgt ggggaaccac atccccctgg atgccctcca gccttggttt 2700

gggctggagt cagggcctgt atacagtatt ttgaatttgt atgccactgg tttgcattgc 2760

tggtcaggaa ctctagtgct ttgcatagcc ctggtttaga aacatgttat agcagttctt 2820

ggtatagagc aaactagaag aaccagcaat cattccactg tcctgccaag gtacacctca 2880

gtactcccct tcccaactga agtggtatga ggctagctct ttccaaaagc attcaagttt 2940

ggcttctgat gtgactcaga atttaggaac cagatgctag atcaaataag ctctgaaaat 3000

ctgaggaaca ttgtaggaaa ggtttgttaa gcatctctta agtgccatga tgagcataac 3060

agccggccgt cgtggctcac gcctgtaatc ccagcacttt gggaggccaa ggtgggagga 3120

tgacaaggtc aggagttcaa gaccagcctg gccaacatgc tgaaacctca cctctactaa 3180

aaatacaaaa attagctggg catggtggca catgcctgta atcccagcta cttgggaggc 3240

tgaggcagga gaatcgcttg aacccgggag gcggaggttg cagtgagcca agacagtgcc 3300

agtgcactcc agcctcggtg acagcgcaag gctccgtctc aataattaaa aaaaaaaaaa 3360

aaaaaaaaaa ggccgggcgc agtggctcaa gcctgtaatc ccagcacttt gggaggctga 3420

ggcgggcaga tcacctgagg tcaggagttt tgagatcagc cttggcaaca cggtgaaacc 3480

ccatctctac taaaaatata aaattagcca agcatgctgg cacatgcctg taatcccagc 3540

tactcgggag gctgaggtac gagaatcgct tgaacctggg aggcagagga tgcagtgagc 3600

cgagatcacg ccattgcact ccagcctggg ggacaagagt gaatctgtgt ctcaccaaaa 3660

aaaaaaagaa aaagaaagat gcttaacaaa ggttaccata agccacaaat tcataaccac 3720

ttatccttcc agtttcaagt agaatatatt cataacctca ataaagttct ccctgctccc 3780

aaa 3783

<210> SEQ ID NO 8
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 8

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
50 55 60

-continued

```
Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65                  70                  75                  80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
                85                  90                  95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
            100                 105                 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
        115                 120                 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
    130                 135                 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145                 150                 155                 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
                165                 170                 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
            180                 185                 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
        195                 200                 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
    210                 215                 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225                 230                 235                 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
                245                 250                 255

Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
            260                 265                 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
        275                 280                 285

Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
    290                 295                 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305                 310                 315                 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
                325                 330                 335

Glu Lys Ile
```

<210> SEQ ID NO 9
<211> LENGTH: 1825
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 9

```
acacatgctg catacacaca gaaacactgc aaatccactg cctccttccc tcctccctac     60

ccttccttct ctcagcattt ctatccccgc ctcctcctct tacccaaatt ttccagccga    120

tcactggagc tgacttccgc aatcccgatg gaataaatct agcacccctg atggtgtgcc    180

cacactttgc tgccgaaacg aagccagaca acagatttcc atcagcagga tgtgggggct    240

caaggttctg ctgctacctg tggtgagctt tgctctgtac cctgaggaga tactggacac    300

ccactgggag ctatggaaga agacccacag gaagcaatat aacaacaagg tggatgaaat    360

ctctcggcgt ttaatttggg aaaaaaacct gaagtatatt tccatccata accttgaggc    420

ttctcttggt gtccatacat atgaactggc tatgaaccac ctgggggaca tgaccagtga    480
```

-continued

agaggtggtt cagaagatga ctggactcaa agtacccctg tctcattccc gcagtaatga 540

cacccтttat atcccagaat gggaaggtag agccccagac tctgtcgact atcgaaagaa 600

aggatatgtt actcctgtca aaaatcaggg tcagtgtggt tcctgttggg cttttagctc 660

tgtgggtgcc ctggagggcc aactcaagaa gaaaactggc aaactcttaa atctgagtcc 720

ccagaaccta gtggattgtg tgtctgagaa tgatggctgt ggaggggggct acatgaccaa 780

tgccttccaa tatgtgcaga agaaccgggg tattgactct gaagatgcct acccatatgt 840

gggacaggaa gagagttgta tgtacaaccc aacaggcaag gcagctaaat gcagagggta 900

cagagagatc cccgagggga atgagaaagc cctgaagagg gcagtggccc gagtgggacc 960

tgtctctgtg gccattgatg caagcctgac ctccttccag ttttacagca aaggtgtgta 1020

ttatgatgaa agctgcaata gcgataatct gaaccatgcg gttttggcag tgggatatgg 1080

aatccagaag ggaaacaagc actggataat taaaaacagc tggggagaaa actggggaaa 1140

caaaggatat atcctcatgg ctcgaaataa gaacaacgcc tgtggcattg ccaacctggc 1200

cagcttcccc aagatgtgac tccagccagc caaatccatc ctgctcttcc atttcttcca 1260

cgatggtgca gtgtaacgat gcactttgga agggagttgg tgtgctattt ttgaagcaga 1320

tgtggtgata ctgagattgt ctgttcagtt tccccatttg tttgtgcttc aaatgatcct 1380

tcctactttg cttctctcca cccatgacct ttttcactgt ggccatcagg actttccctg 1440

acagctgtgt actcttaggc taagagatgt gactacagcc tgcccctgac tgtgttgtcc 1500

cagggctgat gctgtacagg tacaggctgg agattttcac ataggttaga ttctcattca 1560

cgggactagt tagctttaag caccctagag gactagggta atctgacttc tcacttccta 1620

agttcccttc tatatcctca aggtagaaat gtctatgttt tctactccaa ttcataaatc 1680

tattcataag tctttggtac aagtttacat gataaaaaga aatgtgattt gtcttccctt 1740

ctttgcactt ttgaaataaa gtatttatct cctgtctaca gtttaataaa tagcatctag 1800

tacacattca aaaaaaaaaa aaaaa 1825

<210> SEQ ID NO 10
<211> LENGTH: 329
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 10

Met Trp Gly Leu Lys Val Leu Leu Leu Pro Val Val Ser Phe Ala Leu
1 5 10 15

Tyr Pro Glu Glu Ile Leu Asp Thr His Trp Glu Leu Trp Lys Lys Thr
20 25 30

His Arg Lys Gln Tyr Asn Asn Lys Val Asp Glu Ile Ser Arg Arg Leu
35 40 45

Ile Trp Glu Lys Asn Leu Lys Tyr Ile Ser Ile His Asn Leu Glu Ala
50 55 60

Ser Leu Gly Val His Thr Tyr Glu Leu Ala Met Asn His Leu Gly Asp
65 70 75 80

Met Thr Ser Glu Glu Val Val Gln Lys Met Thr Gly Leu Lys Val Pro
85 90 95

Leu Ser His Ser Arg Ser Asn Asp Thr Leu Tyr Ile Pro Glu Trp Glu
100 105 110

-continued

```
Gly Arg Ala Pro Asp Ser Val Asp Tyr Arg Lys Lys Gly Tyr Val Thr
        115                 120                 125

Pro Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ser
    130                 135                 140

Val Gly Ala Leu Glu Gly Gln Leu Lys Lys Lys Thr Gly Lys Leu Leu
145                 150                 155                 160

Asn Leu Ser Pro Gln Asn Leu Val Asp Cys Val Ser Glu Asn Asp Gly
                165                 170                 175

Cys Gly Gly Gly Tyr Met Thr Asn Ala Phe Gln Tyr Val Gln Lys Asn
            180                 185                 190

Arg Gly Ile Asp Ser Glu Asp Ala Tyr Pro Tyr Val Gly Gln Glu Glu
        195                 200                 205

Ser Cys Met Tyr Asn Pro Thr Gly Lys Ala Ala Lys Cys Arg Gly Tyr
    210                 215                 220

Arg Glu Ile Pro Glu Gly Asn Glu Lys Ala Leu Lys Arg Ala Val Ala
225                 230                 235                 240

Arg Val Gly Pro Val Ser Val Ala Ile Asp Ala Ser Leu Thr Ser Phe
                245                 250                 255

Gln Phe Tyr Ser Lys Gly Val Tyr Tyr Asp Glu Ser Cys Asn Ser Asp
            260                 265                 270

Asn Leu Asn His Ala Val Leu Ala Val Gly Tyr Gly Ile Gln Lys Gly
        275                 280                 285

Asn Lys His Trp Ile Ile Lys Asn Ser Trp Gly Glu Asn Trp Gly Asn
    290                 295                 300

Lys Gly Tyr Ile Leu Met Ala Arg Asn Lys Asn Asn Ala Cys Gly Ile
305                 310                 315                 320

Ala Asn Leu Ala Ser Phe Pro Lys Met
                325
```

<210> SEQ ID NO 11
<211> LENGTH: 1730
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 11

```
ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc     60

agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac    120

cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag    180

gacgcggtcg agtaggtgtg caccagccct ggcaacgaga gcgtctaccc cgaactctgc    240

tggccttgag gtggggaagc cggggagggc agttgaggac cccgcggagg cgcgtgactg    300

gttgagcggg caggccagcc tccgagccgg gtggacacag gttttaaaac atgaatccta    360

cactcatcct tgctgccttt tgcctgggaa ttgcctcagc tactctaaca tttgatcaca    420

gtttagaggc acagtggacc aagtggaagg cgatgcacaa cagattatac ggcatgaatg    480

aagaaggatg gaggagagca gtgtgggaga agaacatgaa gatgattgaa ctgcacaatc    540

aggaatacag ggaagggaaa cacagcttca caatggccat gaacgccttt ggagacatga    600

ccagtgaaga attcaggcag gtgatgaatg gctttcaaaa ccgtaagccc aggaagggga    660

aagtgttcca ggaacctctg ttttatgagg cccccagatc tgtggattgg agagagaaag    720

gctacgtgac tcctgtgaag aatcagggtc agtgtggttc ttgttgggct tttagtgcta    780
```

-continued

ctggtgctct tgaaggacag atgttccgga aaactgggag gcttatctca ctgagtgagc 840

agaatctggt agactgctct gggcctcaag gcaatgaagg ctgcaatggt ggcctaatgg 900

attatgcttt ccagtatgtt caggataatg gaggcctgga ctctgaggaa tcctatccat 960

atgaggcaac agaagaatcc tgtaagtaca atcccaagta ttctgttgct aatgacaccg 1020

gctttgtgga catccctaag caggagaagg ccctgatgaa ggcagttgca actgtggggc 1080

ccatttctgt tgctattgat gcaggtcatg agtccttcct gttctataaa gaaggcattt 1140

attttgagcc agactgtagc agtgaagaca tggatcatgg tgtgctggtg gttggctacg 1200

gatttgaaag cacagaatca gataacaata aatattggct ggtgaagaac agctggggtg 1260

aagaatgggg catgggtggc tacgtaaaga tggccaaaga ccggagaaac cattgtggaa 1320

ttgcctcagc agccagctac cccactgtgt gagctggtgg acggtgatga ggaaggactt 1380

gactggggat ggcgcatgca tgggaggaat tcatcttcag tctaccagcc cccgctgtgt 1440

cggatacaca ctcgaatcat tgaagatccg agtgtgattt gaattctgtg atattttcac 1500

actggtaaat gttacctcta ttttaattac tgctataaat aggtttatat tattgattca 1560

cttactgact ttgcattttc gtttttaaaa ggatgtataa atttttacct gtttaaataa 1620

aatttaattt caaatgtagt ggtggggctt ctttctattt ttgatgcact gaatttttgt 1680

gtaataaaga acataattgg gctctaagcc ataaaaaaaa aaaaaaaaaa 1730

<210> SEQ ID NO 12
<211> LENGTH: 333
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 12

Met Asn Pro Thr Leu Ile Leu Ala Ala Phe Cys Leu Gly Ile Ala Ser
1 5 10 15

Ala Thr Leu Thr Phe Asp His Ser Leu Glu Ala Gln Trp Thr Lys Trp
20 25 30

Lys Ala Met His Asn Arg Leu Tyr Gly Met Asn Glu Glu Gly Trp Arg
35 40 45

Arg Ala Val Trp Glu Lys Asn Met Lys Met Ile Glu Leu His Asn Gln
50 55 60

Glu Tyr Arg Glu Gly Lys His Ser Phe Thr Met Ala Met Asn Ala Phe
65 70 75 80

Gly Asp Met Thr Ser Glu Glu Phe Arg Gln Val Met Asn Gly Phe Gln
85 90 95

Asn Arg Lys Pro Arg Lys Gly Lys Val Phe Gln Glu Pro Leu Phe Tyr
100 105 110

Glu Ala Pro Arg Ser Val Asp Trp Arg Glu Lys Gly Tyr Val Thr Pro
115 120 125

Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ala Thr
130 135 140

Gly Ala Leu Glu Gly Gln Met Phe Arg Lys Thr Gly Arg Leu Ile Ser
145 150 155 160

Leu Ser Glu Gln Asn Leu Val Asp Cys Ser Gly Pro Gln Gly Asn Glu
165 170 175

Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Gln Tyr Val Gln Asp
180 185 190

-continued

```
Asn Gly Gly Leu Asp Ser Glu Glu Ser Tyr Pro Tyr Glu Ala Thr Glu
        195                 200                 205

Glu Ser Cys Lys Tyr Asn Pro Lys Tyr Ser Val Ala Asn Asp Thr Gly
    210                 215                 220

Phe Val Asp Ile Pro Lys Gln Glu Lys Ala Leu Met Lys Ala Val Ala
225                 230                 235                 240

Thr Val Gly Pro Ile Ser Val Ala Ile Asp Ala Gly His Glu Ser Phe
                245                 250                 255

Leu Phe Tyr Lys Glu Gly Ile Tyr Phe Glu Pro Asp Cys Ser Ser Glu
            260                 265                 270

Asp Met Asp His Gly Val Leu Val Val Gly Tyr Gly Phe Glu Ser Thr
        275                 280                 285

Glu Ser Asp Asn Asn Lys Tyr Trp Leu Val Lys Asn Ser Trp Gly Glu
    290                 295                 300

Glu Trp Gly Met Gly Gly Tyr Val Lys Met Ala Lys Asp Arg Arg Asn
305                 310                 315                 320

His Cys Gly Ile Ala Ser Ala Ala Ser Tyr Pro Thr Val
                325                 330


<210> SEQ ID NO 13
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(3)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid

<400> SEQUENCE: 13

Asp Xaa Xaa Leu Leu
1               5


<210> SEQ ID NO 14
<211> LENGTH: 8
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (3)..(5)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid

<400> SEQUENCE: 14

Asp Glu Xaa Xaa Xaa Leu Leu Ile
1               5


<210> SEQ ID NO 15
<211> LENGTH: 4
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide
<220> FEATURE:
<221> NAME/KEY: misc_feature
<222> LOCATION: (2)..(3)
<223> OTHER INFORMATION: Xaa can be any naturally occurring amino acid
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (4)..(4)
<223> OTHER INFORMATION: Xaa may be an amino acid with a bulky
```

-continued

hydrophobic side chain, such as Ile, Phe, Leu, Val, and Met

<400> SEQUENCE: 15

Tyr Xaa Xaa Xaa
1

<210> SEQ ID NO 16
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 16

Ser Phe His Asp Asp Ser Asp Glu Asp Leu Leu
1 5 10

<210> SEQ ID NO 17
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 17

Glu Glu Ser Glu Glu Arg Asp Asp His Leu Leu
1 5 10

<210> SEQ ID NO 18
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 18

Gly Tyr His Asp Asp Ser Asp Glu Asp Leu Leu
1 5 10

<210> SEQ ID NO 19
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 19

Ile Thr Gly Phe Ser Asp Asp Val Pro Met Val
1 5 10

<210> SEQ ID NO 20
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 20

Ala Ser Val Ser Leu Leu Asp Asp Glu Leu Met
1 5 10

<210> SEQ ID NO 21
<211> LENGTH: 11
<212> TYPE: PRT

-continued

<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 21

Ala Ser Ser Gly Leu Asp Asp Leu Asp Leu Leu
1 5 10

<210> SEQ ID NO 22
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 22

Val Gln Asn Pro Ser Ala Asp Arg Asn Leu Leu
1 5 10

<210> SEQ ID NO 23
<211> LENGTH: 11
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 23

Asn Ala Leu Ser Trp Leu Asp Glu Glu Leu Leu
1 5 10

<210> SEQ ID NO 24
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 24

Asp Glu Arg Ala Pro Leu Ile
1 5

<210> SEQ ID NO 25
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 25

Thr Glu Arg Glu Arg Leu Leu
1 5

<210> SEQ ID NO 26
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 26

Ser Glu Thr Glu Arg Leu Leu
1 5

<210> SEQ ID NO 27

-continued

<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 27

Thr Asp Arg Thr Pro Leu Leu
1 5

<210> SEQ ID NO 28
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 28

Glu Glu Thr Gln Pro Leu Leu
1 5

<210> SEQ ID NO 29
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 29

Asp Asp Gln Arg Asp Leu Ile
1 5

<210> SEQ ID NO 30
<211> LENGTH: 7
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 30

Asn Glu Gln Leu Pro Met Leu
1 5

<210> SEQ ID NO 31
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 31

Gly Tyr Gln Thr Ile
1 5

<210> SEQ ID NO 32
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 32

Gly Tyr Glu Gln Phe
1 5

-continued

<210> SEQ ID NO 33
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 33

Gly Tyr Gln Thr Leu
1 5

<210> SEQ ID NO 34
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 34

Gly Tyr Gln Ser Val
1 5

<210> SEQ ID NO 35
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 35

Gly Tyr Glu Val Met
1 5

<210> SEQ ID NO 36
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 36

Ala Tyr Gln Ala Leu
1 5

<210> SEQ ID NO 37
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 37

Asn Tyr His Thr Leu
1 5

<210> SEQ ID NO 38
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 38

Gly Tyr Gln Arg Ile

-continued

1 5

<210> SEQ ID NO 39
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 39

Gly Tyr Asp Gln Leu
1 5

<210> SEQ ID NO 40
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 40

Gly Tyr Lys Glu Ile
1 5

<210> SEQ ID NO 41
<211> LENGTH: 5
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: signal peptide

<400> SEQUENCE: 41

Gly Tyr Arg His Val
1 5

<210> SEQ ID NO 42
<211> LENGTH: 2300
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 42

agagtgcacc cgaatccacg ggctcggagg cagcagccat ctctcggcca tagggcaggc 60

cagctggcgc cgggggctat tttgggcggc gggcaatgat ggtgaccgca aggcgacctt 120

gtaaggcatt tccccctga ctcccttccc cgagcctctg cccgggggtc ctagcgccgc 180

tttctcagcc atcccgccta caacttagcc gtccacaaca ggatcatctg atcgcgtgcg 240

cccgggctac gatctgcgag gcccgcggac cttgacccgg cattgaccgc caccgccccc 300

caggtccgta gggaccaaag aaggggcggg aggaagactg tcacgtggcg ccggagttca 360

cgtgactcgt acacatgact tccagtcccc gggcgcctcc tggagagcaa ggacgcgggg 420

gagcagaggt gagctggcac cggaggctgg aggggatccc cgagcccggg atcgatgatc 480

cgagccgcgc cgccgccgct gttcctgctg ctgctgctgc tgctgctgct agtgtcctgg 540

gcgtcccgag gcgaggcagc ccccgaccag gacgagatcc agcgcctccc cgggctggcc 600

aagcagccgt ctttccgcca gtactccggc tacctcaaag gctccggctc caagcacctc 660

cactactggt ttgtggagtc ccagaaggat cccgagaaca gccctgtggt gctttggctc 720

aatgggggtc ccggctgcag ctcactagat gggctcctca cagagcatgg ccccttcctg 780

gtccagccag atggtgtcac cctggagtac aacccctatt cttggaatct gattgccaat 840

-continued

gtgttatacc tggagtcccc agctggggtg ggcttctcct actccgatga caagttttat 900

gcaactaatg acactgaggt cgcccagagc aattttgagg cccttcaaga tttcttccgc 960

ctctttccgg agtacaagaa caacaaactt ttcctgaccg gggagagcta tgctggcatc 1020

tacatcccca ccctggccgt gctggtcatg caggatccca gcatgaacct tcaggggctg 1080

gctgtgggca atggactctc ctcctatgag cagaatgaca actccctggt ctactttgcc 1140

tactaccatg gccttctggg gaacaggctt tggtcttctc tccagaccca ctgctgctct 1200

caaaacaagt gtaacttcta tgacaacaaa gacctggaat gcgtgaccaa tcttcaggaa 1260

gtggcccgca tcgtgggcaa ctctggcctc aacatctaca atctctatgc cccgtgtgct 1320

ggaggggtgc ccagccattt taggtatgag aaggacactg ttgtggtcca ggatttgggc 1380

aacatcttca ctcgcctgcc actcaagcgg atgtggcatc aggcactgct gcgctcaggg 1440

gataaagtgc gcatggaccc cccctgcacc aacacaacag ctgcttccac ctacctcaac 1500

aacccgtacg tgcggaaggc cctcaacatc ccggagcagc tgccacaatg ggacatgtgc 1560

aactttctgg taaacttaca gtaccgccgt ctctaccgaa gcatgaactc ccagtatctg 1620

aagctgctta gctcacagaa ataccagatc ctattatata atggagatgt agacatggcc 1680

tgcaatttca tgggggatga gtggtttgtg gattccctca accagaagat ggaggtgcag 1740

cgccggccct ggttagtgaa gtacggggac agcggggagc agattgccgg cttcgtgaag 1800

gagttctccc acatcgcctt tctcacgatc aagggcgccg gccacatggt tcccaccgac 1860

aagcccctcg ctgccttcac catgttctcc cgcttcctga acaagcagcc atactgatga 1920

ccacagcaac cagctccacg gcctgatgca gcccctccca gcctctcccg ctaggagagt 1980

cctcttctaa gcaaagtgcc cctgcaggcc gggttctgcc gccaggactg ccccttcccc 2040

agagccctgt acatcccaga ctgggcccag ggtctcccat agacagcctg ggggcaagtt 2100

agcactttat tcccgcagca gttcctgaat ggggtggcct ggccccttct ctgcttaaag 2160

aatgcccttt atgatgcact gattccatcc caggaaccca acagagctca ggacagccca 2220

cagggaggtg gtggacggac tgtaattgat agattgatta tggaattaaa ttgggtacag 2280

cttcaaaaaa aaaaaaaaaa 2300

<210> SEQ ID NO 43
<211> LENGTH: 480
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 43

Met Ile Arg Ala Ala Pro Pro Pro Leu Phe Leu Leu Leu Leu Leu Leu
1 5 10 15

Leu Leu Leu Val Ser Trp Ala Ser Arg Gly Glu Ala Ala Pro Asp Gln
20 25 30

Asp Glu Ile Gln Arg Leu Pro Gly Leu Ala Lys Gln Pro Ser Phe Arg
35 40 45

Gln Tyr Ser Gly Tyr Leu Lys Gly Ser Gly Ser Lys His Leu His Tyr
50 55 60

Trp Phe Val Glu Ser Gln Lys Asp Pro Glu Asn Ser Pro Val Val Leu
65 70 75 80

Trp Leu Asn Gly Gly Pro Gly Cys Ser Ser Leu Asp Gly Leu Leu Thr
85 90 95

-continued

```
Glu His Gly Pro Phe Leu Val Gln Pro Asp Gly Val Thr Leu Glu Tyr
            100                 105                 110

Asn Pro Tyr Ser Trp Asn Leu Ile Ala Asn Val Leu Tyr Leu Glu Ser
        115                 120                 125

Pro Ala Gly Val Gly Phe Ser Tyr Ser Asp Asp Lys Phe Tyr Ala Thr
    130                 135                 140

Asn Asp Thr Glu Val Ala Gln Ser Asn Phe Glu Ala Leu Gln Asp Phe
145                 150                 155                 160

Phe Arg Leu Phe Pro Glu Tyr Lys Asn Asn Lys Leu Phe Leu Thr Gly
                165                 170                 175

Glu Ser Tyr Ala Gly Ile Tyr Ile Pro Thr Leu Ala Val Leu Val Met
            180                 185                 190

Gln Asp Pro Ser Met Asn Leu Gln Gly Leu Ala Val Gly Asn Gly Leu
        195                 200                 205

Ser Ser Tyr Glu Gln Asn Asp Asn Ser Leu Val Tyr Phe Ala Tyr Tyr
    210                 215                 220

His Gly Leu Leu Gly Asn Arg Leu Trp Ser Ser Leu Gln Thr His Cys
225                 230                 235                 240

Cys Ser Gln Asn Lys Cys Asn Phe Tyr Asp Asn Lys Asp Leu Glu Cys
                245                 250                 255

Val Thr Asn Leu Gln Glu Val Ala Arg Ile Val Gly Asn Ser Gly Leu
            260                 265                 270

Asn Ile Tyr Asn Leu Tyr Ala Pro Cys Ala Gly Gly Val Pro Ser His
        275                 280                 285

Phe Arg Tyr Glu Lys Asp Thr Val Val Val Gln Asp Leu Gly Asn Ile
    290                 295                 300

Phe Thr Arg Leu Pro Leu Lys Arg Met Trp His Gln Ala Leu Leu Arg
305                 310                 315                 320

Ser Gly Asp Lys Val Arg Met Asp Pro Pro Cys Thr Asn Thr Thr Ala
                325                 330                 335

Ala Ser Thr Tyr Leu Asn Asn Pro Tyr Val Arg Lys Ala Leu Asn Ile
            340                 345                 350

Pro Glu Gln Leu Pro Gln Trp Asp Met Cys Asn Phe Leu Val Asn Leu
        355                 360                 365

Gln Tyr Arg Arg Leu Tyr Arg Ser Met Asn Ser Gln Tyr Leu Lys Leu
    370                 375                 380

Leu Ser Ser Gln Lys Tyr Gln Ile Leu Leu Tyr Asn Gly Asp Val Asp
385                 390                 395                 400

Met Ala Cys Asn Phe Met Gly Asp Glu Trp Phe Val Asp Ser Leu Asn
                405                 410                 415

Gln Lys Met Glu Val Gln Arg Arg Pro Trp Leu Val Lys Tyr Gly Asp
            420                 425                 430

Ser Gly Glu Gln Ile Ala Gly Phe Val Lys Glu Phe Ser His Ile Ala
        435                 440                 445

Phe Leu Thr Ile Lys Gly Ala Gly His Met Val Pro Thr Asp Lys Pro
    450                 455                 460

Leu Ala Ala Phe Thr Met Phe Ser Arg Phe Leu Asn Lys Gln Pro Tyr
465                 470                 475                 480
```

<210> SEQ ID NO 44
<211> LENGTH: 2208

-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 44

agagtgcacc cgaatccacg ggctcggagg cagcagccat ctctcggcca tagggcaggc 60

cagctggcgc cgggggctat tttgggcggc gggcaatgat ggtgaccgca aggcgacctt 120

gtaaggcatt tccccccctga ctcccttccc cgagcctctg cccgggggtc ctagcgccgc 180

tttctcagcc atcccgccta caacttagcc gtccacaaca ggatcatctg atcgcgtgcg 240

cccgggctac gatctgcgag gcccgcggac cttgacccgg cattgaccgc caccgccccc 300

caggtccgta gggaccaaag aaggggcggg aggaagactg tcacgtggcg ccggagttca 360

cgtgactcgt acacatgact tccagtcccc gggcgcctcc tggagagcaa ggacgcgggg 420

gagcagagat gatccgagcc gcgccgccgc cgctgttcct gctgctgctg ctgctgctgc 480

tgctagtgtc ctgggcgtcc cgaggcgagg cagcccccga ccaggacgag atccagcgcc 540

tccccgggct ggccaagcag ccgtctttcc gccagtactc cggctacctc aaaggctccg 600

gctccaagca cctccactac tggtttgtgg agtcccagaa ggatcccgag aacagccctg 660

tggtgctttg gctcaatggg ggtcccggct gcagctcact agatgggctc ctcacagagc 720

atggcccctt cctgattgcc aatgtgttat acctggagtc cccagctggg gtgggcttct 780

cctactccga tgacaagttt tatgcaacta atgacactga ggtcgcccag agcaattttg 840

aggcccttca agatttcttc cgcctctttc cggagtacaa gaacaacaaa cttttcctga 900

ccggggagag ctatgctggc atctacatcc ccaccctggc cgtgctggtc atgcaggatc 960

ccagcatgaa ccttcagggg ctggctgtgg gcaatggact ctcctcctat gagcagaatg 1020

acaactccct ggtctacttt gcctactacc atggccttct ggggaacagg ctttggtctt 1080

ctctccagac ccactgctgc tctcaaaaca agtgtaactt ctatgacaac aaagacctgg 1140

aatgcgtgac caatcttcag gaagtggccc gcatcgtggg caactctggc ctcaacatct 1200

acaatctcta tgccccgtgt gctggagggg tgcccagcca ttttaggtat gagaaggaca 1260

ctgttgtggt ccaggatttg ggcaacatct tcactcgcct gccactcaag cggatgtggc 1320

atcaggcact gctgcgctca ggggataaag tgcgcatgga ccccccctgc accaacacaa 1380

cagctgcttc cacctacctc aacaacccgt acgtgcggaa ggccctcaac atcccggagc 1440

agctgccaca atgggacatg tgcaactttc tggtaaactt acagtaccgc cgtctctacc 1500

gaagcatgaa ctcccagtat ctgaagctgc ttagctcaca gaaataccag atcctattat 1560

ataatggaga tgtagacatg gcctgcaatt tcatggggga tgagtggttt gtggattccc 1620

tcaaccagaa gatggaggtg cagcgccggc cctggttagt gaagtacggg gacagcgggg 1680

agcagattgc cggcttcgtg aaggagttct cccacatcgc ctttctcacg atcaagggcg 1740

ccggccacat ggttcccacc gacaagcccc tcgctgcctt caccatgttc tcccgcttcc 1800

tgaacaagca gccatactga tgaccacagc aaccagctcc acggcctgat gcagcccctc 1860

ccagcctctc ccgctaggag agtcctcttc taagcaaagt gcccctgcag gccgggttct 1920

gccgccagga ctgccccctt cccagagccc tgtacatccc agactgggcc cagggtctcc 1980

catagacagc ctgggggcaa gttagcactt tattcccgca gcagttcctg aatggggtgg 2040

cctggccccct tctctgctta aagaatgccc tttatgatgc actgattcca tcccaggaac 2100

ccaacagagc tcaggacagc ccacagggag gtggtggacg gactgtaatt gatagattga 2160

-continued

ttatggaatt aaattgggta cagcttcaaa aaaaaaaaaa aaaaaaaa 2208

<210> SEQ ID NO 45
<211> LENGTH: 481
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 45

Met Thr Ser Ser Pro Arg Ala Pro Pro Gly Glu Gln Gly Arg Gly Gly
1 5 10 15

Ala Glu Met Ile Arg Ala Ala Pro Pro Pro Leu Phe Leu Leu Leu Leu
20 25 30

Leu Leu Leu Leu Leu Val Ser Trp Ala Ser Arg Gly Glu Ala Ala Pro
35 40 45

Asp Gln Asp Glu Ile Gln Arg Leu Pro Gly Leu Ala Lys Gln Pro Ser
50 55 60

Phe Arg Gln Tyr Ser Gly Tyr Leu Lys Gly Ser Gly Ser Lys His Leu
65 70 75 80

His Tyr Trp Phe Val Glu Ser Gln Lys Asp Pro Glu Asn Ser Pro Val
85 90 95

Val Leu Trp Leu Asn Gly Gly Pro Gly Cys Ser Ser Leu Asp Gly Leu
100 105 110

Leu Thr Glu His Gly Pro Phe Leu Ile Ala Asn Val Leu Tyr Leu Glu
115 120 125

Ser Pro Ala Gly Val Gly Phe Ser Tyr Ser Asp Asp Lys Phe Tyr Ala
130 135 140

Thr Asn Asp Thr Glu Val Ala Gln Ser Asn Phe Glu Ala Leu Gln Asp
145 150 155 160

Phe Phe Arg Leu Phe Pro Glu Tyr Lys Asn Asn Lys Leu Phe Leu Thr
165 170 175

Gly Glu Ser Tyr Ala Gly Ile Tyr Ile Pro Thr Leu Ala Val Leu Val
180 185 190

Met Gln Asp Pro Ser Met Asn Leu Gln Gly Leu Ala Val Gly Asn Gly
195 200 205

Leu Ser Ser Tyr Glu Gln Asn Asp Asn Ser Leu Val Tyr Phe Ala Tyr
210 215 220

Tyr His Gly Leu Leu Gly Asn Arg Leu Trp Ser Ser Leu Gln Thr His
225 230 235 240

Cys Cys Ser Gln Asn Lys Cys Asn Phe Tyr Asp Asn Lys Asp Leu Glu
245 250 255

Cys Val Thr Asn Leu Gln Glu Val Ala Arg Ile Val Gly Asn Ser Gly
260 265 270

Leu Asn Ile Tyr Asn Leu Tyr Ala Pro Cys Ala Gly Gly Val Pro Ser
275 280 285

His Phe Arg Tyr Glu Lys Asp Thr Val Val Val Gln Asp Leu Gly Asn
290 295 300

Ile Phe Thr Arg Leu Pro Leu Lys Arg Met Trp His Gln Ala Leu Leu
305 310 315 320

Arg Ser Gly Asp Lys Val Arg Met Asp Pro Pro Cys Thr Asn Thr Thr
325 330 335

Ala Ala Ser Thr Tyr Leu Asn Asn Pro Tyr Val Arg Lys Ala Leu Asn
340 345 350

-continued

```
Ile Pro Glu Gln Leu Pro Gln Trp Asp Met Cys Asn Phe Leu Val Asn
        355                 360                 365

Leu Gln Tyr Arg Arg Leu Tyr Arg Ser Met Asn Ser Gln Tyr Leu Lys
    370                 375                 380

Leu Leu Ser Ser Gln Lys Tyr Gln Ile Leu Leu Tyr Asn Gly Asp Val
385                 390                 395                 400

Asp Met Ala Cys Asn Phe Met Gly Asp Glu Trp Phe Val Asp Ser Leu
                405                 410                 415

Asn Gln Lys Met Glu Val Gln Arg Arg Pro Trp Leu Val Lys Tyr Gly
            420                 425                 430

Asp Ser Gly Glu Gln Ile Ala Gly Phe Val Lys Glu Phe Ser His Ile
        435                 440                 445

Ala Phe Leu Thr Ile Lys Gly Ala Gly His Met Val Pro Thr Asp Lys
    450                 455                 460

Pro Leu Ala Ala Phe Thr Met Phe Ser Arg Phe Leu Asn Lys Gln Pro
465                 470                 475                 480

Tyr
```

<210> SEQ ID NO 46
<211> LENGTH: 3945
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 46

```
ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg     60

ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct    120

ctcggctgca gcgctgggct ggtgtgcagt ggtgcgacca cggctcacgg cagcctcagc    180

cacccagatg taagcgatct ggttcccacc tcagcctccc gagtagtgtc ttcaggccta    240

tggagagcag cttgcgtggg ctgggcctgc agtacctggt ttgcatagat gattggcagg    300

tggatctagg atccggcttc caacatgtgg cagctctggg cctccctctg ctgcctgctg    360

gtgttggcca atgcccggag caggccctct ttccatcccc tgtcggatga gctggtcaac    420

tatgtcaaca aacggaatac cacgtggcag gccgggcaca acttctacaa cgtggacatg    480

agctacttga agaggctatg tggtaccttc ctgggtgggc ccaagccacc ccagagagtt    540

atgtttaccg aggacctgaa gctgcctgca agcttcgatg cacgggaaca atggccacag    600

tgtcccacca tcaaagagat cagagaccag ggctcctgtg gctcctgctg ggccttcggg    660

gctgtggaag ccatctctga ccggatctgc atccacacca atgcgcacgt cagcgtggag    720

gtgtcggcgg aggacctgct cacatgctgt ggcagcatgt gtggggacgg ctgtaatggt    780

ggctatcctg ctgaagcttg gaacttctgg acaagaaaag gcctggtttc tggtggcctc    840

tatgaatccc atgtagggtg cagaccgtac tccatccctc cctgtgagca ccacgtcaac    900

ggctcccggc ccccatgcac gggggaggga gataccccca agtgtagcaa gatctgtgag    960

cctggctaca gcccgaccta caaacaggac aagcactacg gatacaattc ctacagcgtc   1020

tccaatagcg agaaggacat catggccgag atctacaaaa acggccccgt ggagggagct   1080

ttctctgtgt attcggactt cctgctctac aagtcaggag tgtaccaaca cgtcaccgga   1140

gagatgatgg gtggccatgc catccgcatc ctgggctggg gagtggagaa tggcacaccc   1200

tactggctgg ttgccaactc ctggaacact gactggggtg acaatggctt ctttaaaata   1260
```

-continued

ctcagaggac aggatcactg tggaatcgaa tcagaagtgg tggctggaat tccacgcacc 1320

gatcagtact gggaaaagat ctaatctgcc gtgggcctgt cgtgccagtc ctgggggcga 1380

gatcggggta gaaatgcatt ttattcttta agttcacgta agatacaagt ttcagacagg 1440

gtctgaagga ctggattggc caaacatcag acctgtcttc caaggagacc aagtcctggc 1500

tacatcccag cctgtggtta cagtgcagac aggccatgtg agccaccgct gccagcacag 1560

agcgtccttc cccctgtaga ctagtgccgt agggagtacc tgctgcccca gctgactgtg 1620

gccccctccg tgatccatcc atctccaggg agcaagacag agacgcagga atggaaagcg 1680

gagttcctaa caggatgaaa gttcccccat cagttccccc agtacctcca agcaagtagc 1740

tttccacatt tgtcacagaa atcagaggag agacggtgtt gggagccctt tggagaacgc 1800

cagtctccca ggccccctgc atctatcgag tttgcaatgt cacaacctct ctgatcttgt 1860

gctcagcatg attctttaat agaagtttta ttttttcgtg cactctgcta atcatgtggg 1920

tgagccagtg gaacagcggg agacctgtgc tagttttaca gattgcctcc ttatgacgcg 1980

gctcaaaagg aaaccaagtg gtcaggagtt gtttctgacc cactgatctc tactaccaca 2040

aggaaaatag tttaggagaa accagctttt actgtttttg aaaaattaca gcttcaccct 2100

gtcaagttaa caaggaatgc ctgtgccaat aaaagttttc tccaacttga agtctactct 2160

gatgggatct cagatccttt gtcactgcct atagacttgt agctgctgtc tctctttgtc 2220

cctgcagaga atcacgtcct ggaactgcat gttcttgcga ctcttgggac ttcatcttaa 2280

cttctcgctg ccccagccat gttttcaacc atggcatccc tcccccaatt agttccctgt 2340

catcctcgtc aaccttctct gtaagtgcct ggtaagcttg cccttgctta agaactcaaa 2400

acatagctgt gctctatttt tttgttgttg ttgtgactga cagagtgaga ttccgtctcc 2460

caggctggag tgcagtggcg ccttctcagc tcactgcaac ctgcagcctc ctagattcaa 2520

gcgattctcc tgcttcagcc ttccgagtag ctgggatgac aggcactcac caatatgcct 2580

gggtaatttt tgtatttta agtacataca ggatttcacc atgttggcca ggctagtttc 2640

aaactcccgg cctcaggtgg tctgcctgcc tcagcctccc aaagtgttgg gattacaggc 2700

gtgagccact gggccctgcc tgtatttttt atcagccaca aatccagcaa caagctgagg 2760

attcagctca taaaacaggc ttggtgtctt ggtgatctca cataaccaag atgctacccc 2820

gtggggaacc acatcccccc ggatgccctc cagccttggt ttgggctgga gtcagggcct 2880

gtatacagta ttttgaattt gtatgccact ggtttgcatt gctggtcagg aactctagtg 2940

ctttgcatag ccctggttta gaaacatgtt atagcagttc ttggtataga gcaaactaga 3000

agaaccagca atcattccac tgtcctgcca aggtacacct cagtactccc cttcccaact 3060

gaagtggtat gaggctagct ctttccaaaa gcattcaagt ttggcttctg atgtgactca 3120

gaatttagga accagatgct agatcaaata agctctgaaa atctgaggaa cattgtagga 3180

aaggtttgtt aagcatctct taagtgccat gatgagcata acagccggcc gtcgtggctc 3240

acgcctgtaa tcccagcact ttgggaggcc aaggtgggag gatgacaagg tcaggagttc 3300

aagaccagcc tggccaacat gctgaaacct cacctctact aaaaatacaa aaattagctg 3360

ggcatggtgg cacatgcctg taatcccagc tacttgggag gctgaggcag gagaatcgct 3420

tgaacccggg aggcggaggt tgcagtgagc caagacagtg ccagtgcact ccagcctcgg 3480

tgacagcgca aggctccgtc tcaataatta aaaaaaaaaa aaaaaaaaaa aaggccgggc 3540

-continued

gcagtggctc aagcctgtaa tcccagcact ttgggaggct gaggcgggca gatcacctga 3600

ggtcaggagt tttgagatca gccttggcaa cacggtgaaa ccccatctct actaaaaata 3660

caaaattagc caagcatgct ggcacatgcc tgtaatccca gctactcggg aggctgaggt 3720

acgagaatcg cttgaacctg ggaggcagag gatgcagtga gccgagatca cgccattgca 3780

ctccagcctg ggggacaaga gtgaatctgt gtctcaccaa aaaaaaaaag aaaaagaaag 3840

atgcttaaca aaggttacca taagccacaa attcataacc acttatcctt ccagtttcaa 3900

gtagaatata ttcataacct caataaagtt ctccctgctc ccaaa 3945

<210> SEQ ID NO 47
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 47

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
50 55 60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65 70 75 80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
85 90 95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
100 105 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
115 120 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
130 135 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145 150 155 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
165 170 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
180 185 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
195 200 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
210 215 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225 230 235 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
245 250 255

Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
260 265 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
275 280 285

-continued

```
Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
    290                 295                 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305                 310                 315                 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
                325                 330                 335

Glu Lys Ile
```

<210> SEQ ID NO 48
<211> LENGTH: 3902
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 48

```
ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg     60

ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct    120

ctcggctgca gcgctgggtg tcttcaggcc tatggagagc agcttgcgtg ggctgggcct    180

gcagtacctg gtttgcatag atgattggca ggtgggcagc acggggaagg acctgtgagt    240

ggccaacctg gttcaggtgg atctaggatc cggcttccaa catgtggcag ctctgggcct    300

ccctctgctg cctgctggtg ttggccaatg cccggagcag gccctctttc catcccctgt    360

cggatgagct ggtcaactat gtcaacaaac ggaataccac gtggcaggcc gggcacaact    420

tctacaacgt ggacatgagc tacttgaaga ggctatgtgg taccttcctg ggtgggccca    480

agccacccca gagagttatg tttaccgagg acctgaagct gcctgcaagc ttcgatgcac    540

gggaacaatg gccacagtgt cccaccatca aagagatcag agaccagggc tcctgtggct    600

cctgctgggc cttcggggct gtggaagcca tctctgaccg gatctgcatc cacaccaatg    660

cgcacgtcag cgtggaggtg tcggcggagg acctgctcac atgctgtggc agcatgtgtg    720

gggacggctg taatggtggc tatcctgctg aagcttggaa cttctggaca agaaaaggcc    780

tggtttctgg tggcctctat gaatcccatg tagggtgcag accgtactcc atccctccct    840

gtgagcacca cgtcaacggc tcccggcccc catgcacggg ggagggagat acccccaagt    900

gtagcaagat ctgtgagcct ggctacagcc cgacctacaa acaggacaag cactacggat    960

acaattccta cagcgtctcc aatagcgaga aggacatcat ggccgagatc tacaaaaacg   1020

gccccgtgga gggagctttc tctgtgtatt cggacttcct gctctacaag tcaggagtgt   1080

accaacacgt caccggagag atgatgggtg gccatgccat ccgcatcctg ggctggggag   1140

tggagaatgg cacaccctac tggctggttg ccaactcctg gaacactgac tggggtgaca   1200

atggcttctt taaaatactc agaggacagg atcactgtgg aatcgaatca gaagtggtgg   1260

ctggaattcc acgcaccgat cagtactggg aaaagatcta atctgccgtg ggcctgtcgt   1320

gccagtcctg ggggcgagat cggggtagaa atgcatttta ttctttaagt tcacgtaaga   1380

tacaagtttc agacagggtc tgaaggactg gattggccaa acatcagacc tgtcttccaa   1440

ggagaccaag tcctggctac atcccagcct gtggttacag tgcagacagg ccatgtgagc   1500

caccgctgcc agcacagagc gtccttcccc ctgtagacta gtgccgtagg gagtacctgc   1560

tgccccagct gactgtggcc ccctccgtga tccatccatc tccagggagc aagacagaga   1620

cgcaggaatg gaaagcggag ttcctaacag gatgaaagtt cccccatcag ttcccccagt   1680
```

-continued

acctccaagc aagtagcttt ccacatttgt cacagaaatc agaggagaga cggtgttggg 1740

agccctttgg agaacgccag tctcccaggc cccctgcatc tatcgagttt gcaatgtcac 1800

aacctctctg atcttgtgct cagcatgatt ctttaataga agttttattt tttcgtgcac 1860

tctgctaatc atgtgggtga gccagtggaa cagcgggaga cctgtgctag ttttacagat 1920

tgcctcctta tgacgcggct caaaaggaaa ccaagtggtc aggagttgtt tctgacccac 1980

tgatctctac taccacaagg aaaatagttt aggagaaacc agcttttact gtttttgaaa 2040

aattacagct tcaccctgtc aagttaacaa ggaatgcctg tgccaataaa agttttctcc 2100

aacttgaagt ctactctgat gggatctcag atcctttgtc actgcctata gacttgtagc 2160

tgctgtctct ctttgtccct gcagagaatc acgtcctgga actgcatgtt cttgcgactc 2220

ttgggacttc atcttaactt ctcgctgccc cagccatgtt ttcaaccatg gcatccctcc 2280

cccaattagt tccctgtcat cctcgtcaac cttctctgta agtgcctggt aagcttgccc 2340

ttgcttaaga actcaaaaca tagctgtgct ctatttttttt gttgttgttg tgactgacag 2400

agtgagattc cgtctcccag gctggagtgc agtggcgcct tctcagctca ctgcaacctg 2460

cagcctccta gattcaagcg attctcctgc ttcagccttc cgagtagctg ggatgacagg 2520

cactcaccaa tatgcctggg taatttttgt atttttaagt acatacagga tttcaccatg 2580

ttggccaggc tagtttcaaa ctcccggcct caggtggtct gcctgcctca gcctcccaaa 2640

gtgttgggat tacaggcgtg agccactggg ccctgcctgt atttttttatc agccacaaat 2700

ccagcaacaa gctgaggatt cagctcataa aacaggcttg gtgtcttggt gatctcacat 2760

aaccaagatg ctaccccgtg gggaaccaca tccccctgga tgccctccag ccttggtttg 2820

ggctggagtc agggcctgta tacagtattt tgaatttgta tgccactggt ttgcattgct 2880

ggtcaggaac tctagtgctt tgcatagccc tggtttagaa acatgttata gcagttcttg 2940

gtatagagca aactagaaga accagcaatc attccactgt cctgccaagg tacacctcag 3000

tactcccctt cccaactgaa gtggtatgag gctagctctt tccaaaagca ttcaagtttg 3060

gcttctgatg tgactcagaa tttaggaacc agatgctaga tcaaataagc tctgaaaatc 3120

tgaggaacat tgtaggaaag gtttgttaag catctcttaa gtgccatgat gagcataaca 3180

gccggccgtc gtggctcacg cctgtaatcc cagcactttg ggaggccaag gtgggaggat 3240

gacaaggtca ggagttcaag accagcctgg ccaacatgct gaaacctcac ctctactaaa 3300

aatacaaaaa ttagctgggc atggtggcac atgcctgtaa tcccagctac ttgggaggct 3360

gaggcaggag aatcgcttga acccgggagg cggaggttgc agtgagccaa gacagtgcca 3420

gtgcactcca gcctcggtga cagcgcaagg ctccgtctca ataattaaaa aaaaaaaaaa 3480

aaaaaaaaag gccgggcgca gtggctcaag cctgtaatcc cagcactttg ggaggctgag 3540

gcgggcagat cacctgaggt caggagtttt gagatcagcc ttggcaacac ggtgaaaccc 3600

catctctact aaaaatacaa aattagccaa gcatgctggc acatgcctgt aatcccagct 3660

actcgggagg ctgaggtacg agaatcgctt gaacctggga ggcagaggat gcagtgagcc 3720

gagatcacgc cattgcactc cagcctgggg gacaagagtg aatctgtgtc tcaccaaaaa 3780

aaaaaagaaa aagaaagatg cttaacaaag gttaccataa gccacaaatt cataaccact 3840

tatccttcca gtttcaagta gaatatattc ataacctcaa taaagttctc cctgctccca 3900

aa 3902

-continued

<210> SEQ ID NO 49
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 49

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
50 55 60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65 70 75 80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
85 90 95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
100 105 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
115 120 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
130 135 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145 150 155 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
165 170 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
180 185 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
195 200 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
210 215 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225 230 235 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
245 250 255

Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
260 265 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
275 280 285

Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
290 295 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305 310 315 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
325 330 335

Glu Lys Ile

<210> SEQ ID NO 50
<211> LENGTH: 3871

-continued

<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 50

ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg 60

ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct 120

ctcggctgca gcgctgggct ggtgtgcagt ggtgcgacca cggctcacgg cagcctcagc 180

cacccagatg taagcgatct ggttcccacc tcagcctccc gagtagtgga tctaggatcc 240

ggcttccaac atgtggcagc tctgggcctc cctctgctgc ctgctggtgt tggccaatgc 300

ccggagcagg ccctctttcc atcccctgtc ggatgagctg gtcaactatg tcaacaaacg 360

gaataccacg tggcaggccg ggcacaactt ctacaacgtg gacatgagct acttgaagag 420

gctatgtggt accttcctgg gtgggcccaa gccaccccag agagttatgt ttaccgagga 480

cctgaagctg cctgcaagct tcgatgcacg ggaacaatgg ccacagtgtc ccaccatcaa 540

agagatcaga gaccagggct cctgtggctc ctgctgggcc ttcggggctg tggaagccat 600

ctctgaccgg atctgcatcc acaccaatgc gcacgtcagc gtggaggtgt cggcggagga 660

cctgctcaca tgctgtggca gcatgtgtgg ggacggctgt aatggtggct atcctgctga 720

agcttggaac ttctggacaa gaaaaggcct ggtttctggt ggcctctatg aatcccatgt 780

agggtgcaga ccgtactcca tccctccctg tgagcaccac gtcaacggct cccggccccc 840

atgcacgggg gagggagata cccccaagtg tagcaagatc tgtgagcctg gctacagccc 900

gacctacaaa caggacaagc actacggata caattcctac agcgtctcca atagcgagaa 960

ggacatcatg gccgagatct acaaaaacgg ccccgtggag ggagctttct ctgtgtattc 1020

ggacttcctg ctctacaagt caggagtgta ccaacacgtc accggagaga tgatgggtgg 1080

ccatgccatc cgcatcctgg gctggggagt ggagaatggc acaccctact ggctggttgc 1140

caactcctgg aacactgact ggggtgacaa tggcttcttt aaaatactca gaggacagga 1200

tcactgtgga atcgaatcag aagtggtggc tggaattcca cgcaccgatc agtactggga 1260

aaagatctaa tctgccgtgg gcctgtcgtg ccagtcctgg gggcgagatc ggggtagaaa 1320

tgcattttat tctttaagtt cacgtaagat acaagtttca gacagggtct gaaggactgg 1380

attggccaaa catcagacct gtcttccaag gagaccaagt cctggctaca tcccagcctg 1440

tggttacagt gcagacaggc catgtgagcc accgctgcca gcacagagcg tccttccccc 1500

tgtagactag tgccgtaggg agtacctgct gccccagctg actgtggccc cctccgtgat 1560

ccatccatct ccagggagca agacagagac gcaggaatgg aaagcggagt tcctaacagg 1620

atgaaagttc ccccatcagt tcccccagta cctccaagca agtagctttc cacatttgtc 1680

acagaaatca gaggagagac ggtgttggga gccctttgga gaacgccagt ctcccaggcc 1740

ccctgcatct atcgagtttg caatgtcaca acctctctga tcttgtgctc agcatgattc 1800

tttaatagaa gttttatttt ttcgtgcact ctgctaatca tgtgggtgag ccagtggaac 1860

agcgggagac ctgtgctagt tttacagatt gcctccttat gacgcggctc aaaaggaaac 1920

caagtggtca ggagttgttt ctgacccact gatctctact accacaagga aaatagttta 1980

ggagaaacca gcttttactg tttttgaaaa attacagctt caccctgtca agttaacaag 2040

gaatgcctgt gccaataaaa gttttctcca acttgaagtc tactctgatg ggatctcaga 2100

tcctttgtca ctgcctatag acttgtagct gctgtctctc tttgtccctg cagagaatca 2160

-continued

cgtcctggaa ctgcatgttc ttgcgactct tgggacttca tcttaacttc tcgctgcccc 2220

agccatgttt tcaaccatgg catccctccc ccaattagtt ccctgtcatc ctcgtcaacc 2280

ttctctgtaa gtgcctggta agcttgccct tgcttaagaa ctcaaaacat agctgtgctc 2340

tatttttttg ttgttgttgt gactgacaga gtgagattcc gtctcccagg ctggagtgca 2400

gtggcgcctt ctcagctcac tgcaacctgc agcctcctag attcaagcga ttctcctgct 2460

tcagccttcc gagtagctgg gatgacaggc actcaccaat atgcctgggt aatttttgta 2520

tttttaagta catacaggat ttcaccatgt tggccaggct agtttcaaac tcccggcctc 2580

aggtggtctg cctgcctcag cctcccaaag tgttgggatt acaggcgtga gccactgggc 2640

cctgcctgta ttttttatca gccacaaatc cagcaacaag ctgaggattc agctcataaa 2700

acaggcttgg tgtcttggtg atctcacata accaagatgc taccccgtgg ggaaccacat 2760

ccccctggat gccctccagc cttggtttgg gctggagtca gggcctgtat acagtatttt 2820

gaatttgtat gccactggtt tgcattgctg gtcaggaact ctagtgcttt gcatagccct 2880

ggtttagaaa catgttatag cagttcttgg tatagagcaa actagaagaa ccagcaatca 2940

ttccactgtc ctgccaaggt acacctcagt actccccttc ccaactgaag tggtatgagg 3000

ctagctcttt ccaaaagcat tcaagtttgg cttctgatgt gactcagaat ttaggaacca 3060

gatgctagat caaataagct ctgaaaatct gaggaacatt gtaggaaagg tttgttaagc 3120

atctcttaag tgccatgatg agcataacag ccggccgtcg tggctcacgc ctgtaatccc 3180

agcactttgg gaggccaagg tgggaggatg acaaggtcag gagttcaaga ccagcctggc 3240

caacatgctg aaacctcacc tctactaaaa atacaaaaat tagctgggca tggtggcaca 3300

tgcctgtaat cccagctact tgggaggctg aggcaggaga atcgcttgaa cccgggaggc 3360

ggaggttgca gtgagccaag acagtgccag tgcactccag cctcggtgac agcgcaaggc 3420

tccgtctcaa taattaaaaa aaaaaaaaaa aaaaaaaagg ccgggcgcag tggctcaagc 3480

ctgtaatccc agcactttgg gaggctgagg cgggcagatc acctgaggtc aggagttttg 3540

agatcagcct tggcaacacg gtgaaacccc atctctacta aaaatacaaa attagccaag 3600

catgctggca catgcctgta atcccagcta ctcgggaggc tgaggtacga gaatcgcttg 3660

aacctgggag gcagaggatg cagtgagccg agatcacgcc attgcactcc agcctggggg 3720

acaagagtga atctgtgtct caccaaaaaa aaaaagaaaa agaaagatgc ttaacaaagg 3780

ttaccataag ccacaaattc ataaccactt atccttccag tttcaagtag aatatattca 3840

taacctcaat aaagttctcc ctgctcccaa a 3871

<210> SEQ ID NO 51
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 51

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

-continued

```
Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
    50                  55                  60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65                  70                  75                  80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
                85                  90                  95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
            100                 105                 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
        115                 120                 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
    130                 135                 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145                 150                 155                 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
                165                 170                 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
            180                 185                 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
        195                 200                 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
    210                 215                 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225                 230                 235                 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
                245                 250                 255

Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
            260                 265                 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
        275                 280                 285

Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
    290                 295                 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305                 310                 315                 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
                325                 330                 335

Glu Lys Ile
```

<210> SEQ ID NO 52
<211> LENGTH: 3857
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 52

```
ggggcggggc cgggagggta cttagggccg gggctggccc aggctacggc ggctgcaggg     60

ctccggcaac cgctccggca acgccaaccg ctccgctgcg cgcaggctgg gctgcaggct    120

ctcggctgca gcgctgggtg tcttcaggcc tatggagagc agcttgcgtg ggctgggcct    180

gcagtacctg gtttgcatag atgattggca ggtggatcta ggatccggct tccaacatgt    240

ggcagctctg ggcctccctc tgctgcctgc tggtgttggc caatgcccgg agcaggccct    300

ctttccatcc cctgtcggat gagctggtca actatgtcaa caaacggaat accacgtggc    360
```

-continued

aggccgggca caacttctac aacgtggaca tgagctactt gaagaggcta tgtggtacct 420

tcctgggtgg gcccaagcca ccccagagag ttatgtttac cgaggacctg aagctgcctg 480

caagcttcga tgcacgggaa caatggccac agtgtcccac catcaaagag atcagagacc 540

agggctcctg tggctcctgc tgggccttcg gggctgtgga agccatctct gaccggatct 600

gcatccacac caatgcgcac gtcagcgtgg aggtgtcggc ggaggacctg ctcacatgct 660

gtggcagcat gtgtggggac ggctgtaatg gtggctatcc tgctgaagct tggaacttct 720

ggacaagaaa aggcctggtt tctggtggcc tctatgaatc ccatgtaggg tgcagaccgt 780

actccatccc tccctgtgag caccacgtca acggctcccg gcccccatgc acgggggagg 840

gagatacccc caagtgtagc aagatctgtg agcctggcta cagcccgacc tacaaacagg 900

acaagcacta cggatacaat tcctacagcg tctccaatag cgagaaggac atcatggccg 960

agatctacaa aaacggcccc gtggagggag ctttctctgt gtattcggac ttcctgctct 1020

acaagtcagg agtgtaccaa cacgtcaccg gagagatgat gggtggccat gccatccgca 1080

tcctgggctg gggagtggag aatggcacac cctactggct ggttgccaac tcctggaaca 1140

ctgactgggg tgacaatggc ttctttaaaa tactcagagg acaggatcac tgtggaatcg 1200

aatcagaagt ggtggctgga attccacgca ccgatcagta ctgggaaaag atctaatctg 1260

ccgtgggcct gtcgtgccag tcctgggggc gagatcgggg tagaaatgca ttttattctt 1320

taagttcacg taagatacaa gtttcagaca gggtctgaag gactggattg gccaaacatc 1380

agacctgtct tccaaggaga ccaagtcctg gctacatccc agcctgtggt tacagtgcag 1440

acaggccatg tgagccaccg ctgccagcac agagcgtcct tccccctgta gactagtgcc 1500

gtagggagta cctgctgccc cagctgactg tggccccctc cgtgatccat ccatctccag 1560

ggagcaagac agagacgcag gaatggaaag cggagttcct aacaggatga aagttccccc 1620

atcagttccc ccagtacctc caagcaagta gctttccaca tttgtcacag aaatcagagg 1680

agagacggtg ttgggagccc tttggagaac gccagtctcc caggcccct gcatctatcg 1740

agtttgcaat gtcacaacct ctctgatctt gtgctcagca tgattcttta atagaagttt 1800

tatttttcg tgcactctgc taatcatgtg ggtgagccag tggaacagcg ggagacctgt 1860

gctagtttta cagattgcct ccttatgacg cggctcaaaa ggaaaccaag tggtcaggag 1920

ttgtttctga cccactgatc tctactacca caaggaaaat agtttaggag aaaccagctt 1980

ttactgtttt tgaaaaatta cagcttcacc ctgtcaagtt aacaaggaat gcctgtgcca 2040

ataaaagttt tctccaactt gaagtctact ctgatgggat ctcagatcct ttgtcactgc 2100

ctatagactt gtagctgctg tctctctttg tccctgcaga gaatcacgtc ctggaactgc 2160

atgttcttgc gactcttggg acttcatctt aacttctcgc tgccccagcc atgttttcaa 2220

ccatggcatc cctcccccaa ttagttccct gtcatcctcg tcaaccttct ctgtaagtgc 2280

ctggtaagct tgcccttgct taagaactca aaacatagct gtgctctatt tttttgttgt 2340

tgttgtgact gacagagtga gattccgtct cccaggctgg agtgcagtgg cgccttctca 2400

gctcactgca acctgcagcc tcctagattc aagcgattct cctgcttcag ccttccgagt 2460

agctgggatg acaggcactc accaatatgc ctgggtaatt tttgtatttt taagtacata 2520

caggatttca ccatgttggc caggctagtt tcaaactccc ggcctcaggt ggtctgcctg 2580

cctcagcctc ccaaagtgtt gggattacag gcgtgagcca ctgggccctg cctgtatttt 2640

-continued

ttatcagcca caaatccagc aacaagctga ggattcagct cataaaacag gcttggtgtc 2700

ttggtgatct cacataacca agatgctacc ccgtggggaa ccacatcccc ctggatgccc 2760

tccagccttg gtttgggctg gagtcagggc ctgtatacag tattttgaat ttgtatgcca 2820

ctggtttgca ttgctggtca ggaactctag tgctttgcat agccctggtt tagaaacatg 2880

ttatagcagt tcttggtata gagcaaacta gaagaaccag caatcattcc actgtcctgc 2940

caaggtacac ctcagtactc cccttcccaa ctgaagtggt atgaggctag ctctttccaa 3000

aagcattcaa gtttggcttc tgatgtgact cagaatttag gaaccagatg ctagatcaaa 3060

taagctctga aaatctgagg aacattgtag gaaaggtttg ttaagcatct cttaagtgcc 3120

atgatgagca taacagccgg ccgtcgtggc tcacgcctgt aatcccagca ctttgggagg 3180

ccaaggtggg aggatgacaa ggtcaggagt tcaagaccag cctggccaac atgctgaaac 3240

ctcacctcta ctaaaaatac aaaaattagc tgggcatggt ggcacatgcc tgtaatccca 3300

gctacttggg aggctgaggc aggagaatcg cttgaacccg ggaggcggag gttgcagtga 3360

gccaagacag tgccagtgca ctccagcctc ggtgacagcg caaggctccg tctcaataat 3420

taaaaaaaaa aaaaaaaaaa aaaaggccgg gcgcagtggc tcaagcctgt aatcccagca 3480

ctttgggagg ctgaggcggg cagatcacct gaggtcagga gttttgagat cagccttggc 3540

aacacggtga aaccccatct ctactaaaaa tacaaaatta gccaagcatg ctggcacatg 3600

cctgtaatcc cagctactcg ggaggctgag gtacgagaat cgcttgaacc tgggaggcag 3660

aggatgcagt gagccgagat cacgccattg cactccagcc tgggggacaa gagtgaatct 3720

gtgtctcacc aaaaaaaaaa agaaaaagaa agatgcttaa caaaggttac cataagccac 3780

aaattcataa ccacttatcc ttccagtttc aagtagaata tattcataac ctcaataaag 3840

ttctccctgc tcccaaa 3857

<210> SEQ ID NO 53
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 53

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
50 55 60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65 70 75 80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
85 90 95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
100 105 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
115 120 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
130 135 140

-continued

```
Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145                 150                 155                 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
                165                 170                 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
            180                 185                 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
        195                 200                 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
    210                 215                 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225                 230                 235                 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
                245                 250                 255

Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
            260                 265                 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
        275                 280                 285

Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
    290                 295                 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305                 310                 315                 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
                325                 330                 335

Glu Lys Ile
```

<210> SEQ ID NO 54
<211> LENGTH: 3982
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 54

```
agggccgggg ctggcccagg ctacggcggc tgcagggctc cggcaaccgc tccggcaacg     60

ccaaccgctc cgctgcgcgc aggctgggct gcaggctctc ggctgcagcg ctgggctggt    120

gtgcagtggt gcgaccacgg ctcacggcag cctcagccac ccagatgtaa gcgatctggt    180

tcccacctca gcctcccgag tagatacttc tgaaaataga aatgatgact ctgggatgca    240

aacgttggct gtcctatgta taaggagatg gcttttcacg ctcccagtga ctgaggaagt    300

ttctcccaga tggcgctgct ctgagcctgg tgcagggtgg atctaggatc cggcttccaa    360

catgtggcag ctctgggcct ccctctgctg cctgctggtg ttggccaatg cccggagcag    420

gccctctttc catcccctgt cggatgagct ggtcaactat gtcaacaaac ggaataccac    480

gtggcaggcc gggcacaact tctacaacgt ggacatgagc tacttgaaga ggctatgtgg    540

taccttcctg ggtgggccca agccacccca gagagttatg tttaccgagg acctgaagct    600

gcctgcaagc ttcgatgcac gggaacaatg gccacagtgt cccaccatca aagagatcag    660

agaccagggc tcctgtggct cctgctgggc cttcggggct gtggaagcca tctctgaccg    720

gatctgcatc cacaccaatg cgcacgtcag cgtggaggtg tcggcggagg acctgctcac    780

atgctgtggc agcatgtgtg gggacggctg taatggtggc tatcctgctg aagcttggaa    840

cttctggaca agaaaaggcc tggtttctgg tggcctctat gaatcccatg tagggtgcag    900
```

-continued

accgtactcc atccctccct gtgagcacca cgtcaacggc tcccggcccc catgcacggg 960

ggagggagat acccccaagt gtagcaagat ctgtgagcct ggctacagcc cgacctacaa 1020

acaggacaag cactacggat acaattccta cagcgtctcc aatagcgaga aggacatcat 1080

ggccgagatc tacaaaaacg gccccgtgga gggagctttc tctgtgtatt cggacttcct 1140

gctctacaag tcaggagtgt accaacacgt caccggagag atgatgggtg gccatgccat 1200

ccgcatcctg ggctggggag tggagaatgg cacaccctac tggctggttg ccaactcctg 1260

gaacactgac tggggtgaca atggcttctt taaaatactc agaggacagg atcactgtgg 1320

aatcgaatca gaagtggtgg ctggaattcc acgcaccgat cagtactggg aaaagatcta 1380

atctgccgtg ggcctgtcgt gccagtcctg ggggcgagat cggggtagaa atgcatttta 1440

ttctttaagt tcacgtaaga tacaagtttc agacagggtc tgaaggactg gattggccaa 1500

acatcagacc tgtcttccaa ggagaccaag tcctggctac atcccagcct gtggttacag 1560

tgcagacagg ccatgtgagc caccgctgcc agcacagagc gtccttcccc ctgtagacta 1620

gtgccgtagg gagtacctgc tgccccagct gactgtggcc ccctccgtga tccatccatc 1680

tccagggagc aagacagaga cgcaggaatg gaaagcggag ttcctaacag gatgaaagtt 1740

cccccatcag ttcccccagt acctccaagc aagtagcttt ccacatttgt cacagaaatc 1800

agaggagaga cggtgttggg agccctttgg agaacgccag tctcccaggc cccctgcatc 1860

tatcgagttt gcaatgtcac aacctctctg atcttgtgct cagcatgatt ctttaataga 1920

agttttattt tttcgtgcac tctgctaatc atgtgggtga gccagtggaa cagcgggaga 1980

cctgtgctag ttttacagat tgcctcctta tgacgcggct caaaaggaaa ccaagtggtc 2040

aggagttgtt tctgacccac tgatctctac taccacaagg aaaatagttt aggagaaacc 2100

agcttttact gtttttgaaa aattacagct tcaccctgtc aagttaacaa ggaatgcctg 2160

tgccaataaa agttttctcc aacttgaagt ctactctgat gggatctcag atcctttgtc 2220

actgcctata gacttgtagc tgctgtctct ctttgtccct gcagagaatc acgtcctgga 2280

actgcatgtt cttgcgactc ttgggacttc atcttaactt ctcgctgccc cagccatgtt 2340

ttcaaccatg gcatccctcc cccaattagt tccctgtcat cctcgtcaac cttctctgta 2400

agtgcctggt aagcttgccc ttgcttaaga actcaaaaca tagctgtgct ctatttttt 2460

gttgttgttg tgactgacag agtgagattc cgtctcccag gctggagtgc agtggcgcct 2520

tctcagctca ctgcaacctg cagcctccta gattcaagcg attctcctgc ttcagccttc 2580

cgagtagctg ggatgacagg cactcaccaa tatgcctggg taatttttgt atttttaagt 2640

acatacagga tttcaccatg ttggccaggc tagtttcaaa ctcccggcct caggtggtct 2700

gcctgcctca gcctcccaaa gtgttgggat tacaggcgtg agccactggg ccctgcctgt 2760

atttttatc agccacaaat ccagcaacaa gctgaggatt cagctcataa aacaggcttg 2820

gtgtcttggt gatctcacat aaccaagatg ctaccccgtg gggaaccaca tccccctgga 2880

tgccctccag ccttggtttg ggctggagtc agggcctgta tacagtattt tgaatttgta 2940

tgccactggt ttgcattgct ggtcaggaac tctagtgctt tgcatagccc tggtttagaa 3000

acatgttata gcagttcttg gtatagagca aactagaaga accagcaatc attccactgt 3060

cctgccaagg tacacctcag tactcccctt cccaactgaa gtggtatgag gctagctctt 3120

tccaaaagca ttcaagtttg gcttctgatg tgactcagaa tttaggaacc agatgctaga 3180

-continued

tcaaataagc tctgaaaatc tgaggaacat tgtaggaaag gtttgttaag catctcttaa 3240

gtgccatgat gagcataaca gccggccgtc gtggctcacg cctgtaatcc cagcactttg 3300

ggaggccaag gtgggaggat gacaaggtca ggagttcaag accagcctgg ccaacatgct 3360

gaaacctcac ctctactaaa aatacaaaaa ttagctgggc atggtggcac atgcctgtaa 3420

tcccagctac ttgggaggct gaggcaggag aatcgcttga acccgggagg cggaggttgc 3480

agtgagccaa gacagtgcca gtgcactcca gcctcggtga cagcgcaagg ctccgtctca 3540

ataattaaaa aaaaaaaaaa aaaaaaaaag gccgggcgca gtggctcaag cctgtaatcc 3600

cagcactttg ggaggctgag gcgggcagat cacctgaggt caggagtttt gagatcagcc 3660

ttggcaacac ggtgaaaccc catctctact aaaaatacaa aattagccaa gcatgctggc 3720

acatgcctgt aatcccagct actcgggagg ctgaggtacg agaatcgctt gaacctggga 3780

ggcagaggat gcagtgagcc gagatcacgc cattgcactc cagcctgggg gacaagagtg 3840

aatctgtgtc tcaccaaaaa aaaaaagaaa aagaaagatg cttaacaaag gttaccataa 3900

gccacaaatt cataaccact tatccttcca gtttcaagta gaatatattc ataacctcaa 3960

taaagttctc cctgctccca aa 3982

<210> SEQ ID NO 55
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 55

Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1 5 10 15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
20 25 30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
35 40 45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
50 55 60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65 70 75 80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
85 90 95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
100 105 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
115 120 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
130 135 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145 150 155 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
165 170 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
180 185 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
195 200 205

-continued

```
Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
    210                 215                 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225                 230                 235                 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
                245                 250                 255

Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
            260                 265                 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
        275                 280                 285

Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
    290                 295                 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305                 310                 315                 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
                325                 330                 335

Glu Lys Ile


<210> SEQ ID NO 56
<211> LENGTH: 4086
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 56

caggaccgcc gagggaggcg cctgcgagga agagctcggc cgggtccgga gactgctgcc     60

tgggaccgcg ctcccagcgc ctgggcctcg gtgtctccgg gccaaactgc cgacataatc    120

gcatctgccg gcatctattt tcggtttatt tccccctcat tgcgaaggat ttgcctggcc    180

aactttctgc gcaagatccc acgcaattcc tgggacccca gaagacaggt cctgttgaag    240

aacaggaatc tggcactggg tgggctgggg aggaagccgc acggtgttaa atccataaac    300

aggaagagaa accagacagc gaaaccaaga ggcgaatggg cgattggatg ccggtgggga    360

gaaggccggg ggcgcaccct gctcctggac tccagtaaag ggaggccggg cagagtccct    420

ggggcgccac ctccccctcg gtggatctag gatccggctt ccaacatgtg gcagctctgg    480

gcctccctct gctgcctgct ggtgttggcc aatgcccgga gcaggccctc tttccatccc    540

ctgtcggatg agctggtcaa ctatgtcaac aaacggaata ccacgtggca ggccgggcac    600

aacttctaca acgtggacat gagctacttg aagaggctat gtggtacctt cctgggtggg    660

cccaagccac cccagagagt tatgtttacc gaggacctga agctgcctgc aagcttcgat    720

gcacgggaac aatggccaca gtgtcccacc atcaaagaga tcagagacca gggctcctgt    780

ggctcctgct gggccttcgg ggctgtggaa gccatctctg accggatctg catccacacc    840

aatgcgcacg tcagcgtgga ggtgtcggcg gaggacctgc tcacatgctg tggcagcatg    900

tgtggggacg gctgtaatgg tggctatcct gctgaagctt ggaacttctg gacaagaaaa    960

ggcctggttt ctggtggcct ctatgaatcc catgtagggt gcagaccgta ctccatccct   1020

ccctgtgagc accacgtcaa cggctcccgg cccccatgca cgggggaggg agataccccc   1080

aagtgtagca agatctgtga gcctggctac agcccgacct acaaacagga caagcactac   1140

ggatacaatt cctacagcgt ctccaatagc gagaaggaca tcatggccga gatctacaaa   1200

aacggccccg tggagggagc tttctctgtg tattcggact tcctgctcta caagtcagga   1260
```

-continued

gtgtaccaac acgtcaccgg agagatgatg ggtggccatg ccatccgcat cctgggctgg 1320

ggagtggaga atggcacacc ctactggctg gttgccaact cctggaacac tgactggggt 1380

gacaatggct tctttaaaat actcagagga caggatcact gtggaatcga atcagaagtg 1440

gtggctggaa ttccacgcac cgatcagtac tgggaaaaga tctaatctgc cgtgggcctg 1500

tcgtgccagt cctgggggcg agatcggggt agaaatgcat tttattcttt aagttcacgt 1560

aagatacaag tttcagacag ggtctgaagg actggattgg ccaaacatca gacctgtctt 1620

ccaaggagac caagtcctgg ctacatccca gcctgtggtt acagtgcaga caggccatgt 1680

gagccaccgc tgccagcaca gagcgtcctt ccccctgtag actagtgccg tagggagtac 1740

ctgctgcccc agctgactgt ggccccctcc gtgatccatc catctccagg gagcaagaca 1800

gagacgcagg aatggaaagc ggagttccta acaggatgaa agttccccca tcagttcccc 1860

cagtacctcc aagcaagtag ctttccacat ttgtcacaga aatcagagga gagacggtgt 1920

tgggagccct ttggagaacg ccagtctccc aggccccctg catctatcga gtttgcaatg 1980

tcacaacctc tctgatcttg tgctcagcat gattctttaa tagaagtttt atttttcgt 2040

gcactctgct aatcatgtgg gtgagccagt ggaacagcgg gagacctgtg ctagttttac 2100

agattgcctc cttatgacgc ggctcaaaag gaaaccaagt ggtcaggagt tgtttctgac 2160

ccactgatct ctactaccac aaggaaaata gtttaggaga aaccagcttt tactgttttt 2220

gaaaaattac agcttcaccc tgtcaagtta acaaggaatg cctgtgccaa taaaagtttt 2280

ctccaacttg aagtctactc tgatgggatc tcagatcctt tgtcactgcc tatagacttg 2340

tagctgctgt ctctctttgt ccctgcagag aatcacgtcc tggaactgca tgttcttgcg 2400

actcttggga cttcatctta acttctcgct gccccagcca tgttttcaac catggcatcc 2460

ctcccccaat tagttccctg tcatcctcgt caaccttctc tgtaagtgcc tggtaagctt 2520

gcccttgctt aagaactcaa aacatagctg tgctctattt ttttgttgtt gttgtgactg 2580

acagagtgag attccgtctc ccaggctgga gtgcagtggc gccttctcag ctcactgcaa 2640

cctgcagcct cctagattca agcgattctc ctgcttcagc cttccgagta gctgggatga 2700

caggcactca ccaatatgcc tgggtaattt ttgtattttt aagtacatac aggatttcac 2760

catgttggcc aggctagttt caaactcccg gcctcaggtg gtctgcctgc ctcagcctcc 2820

caaagtgttg ggattacagg cgtgagccac tgggccctgc ctgtattttt tatcagccac 2880

aaatccagca acaagctgag gattcagctc ataaaacagg cttggtgtct tggtgatctc 2940

acataaccaa gatgctaccc cgtggggaac cacatccccc tggatgccct ccagccttgg 3000

tttgggctgg agtcagggcc tgtatacagt attttgaatt tgtatgccac tggtttgcat 3060

tgctggtcag gaactctagt gctttgcata gccctggttt agaaacatgt tatagcagtt 3120

cttggtatag agcaaactag aagaaccagc aatcattcca ctgtcctgcc aaggtacacc 3180

tcagtactcc ccttcccaac tgaagtggta tgaggctagc tctttccaaa agcattcaag 3240

tttggcttct gatgtgactc agaatttagg aaccagatgc tagatcaaat aagctctgaa 3300

aatctgagga acattgtagg aaaggtttgt taagcatctc ttaagtgcca tgatgagcat 3360

aacagccggc cgtcgtggct cacgcctgta atcccagcac tttgggaggc caaggtggga 3420

ggatgacaag gtcaggagtt caagaccagc ctggccaaca tgctgaaacc tcacctctac 3480

taaaaataca aaaattagct gggcatggtg gcacatgcct gtaatcccag ctacttggga 3540

-continued

```
ggctgaggca ggagaatcgc ttgaacccgg gaggcggagg ttgcagtgag ccaagacagt   3600

gccagtgcac tccagcctcg gtgacagcgc aaggctccgt ctcaataatt aaaaaaaaaa   3660

aaaaaaaaaa aaaggccggg cgcagtggct caagcctgta atcccagcac tttgggaggc   3720

tgaggcgggc agatcacctg aggtcaggag ttttgagatc agccttggca acacggtgaa   3780

accccatctc tactaaaaat acaaaattag ccaagcatgc tggcacatgc ctgtaatccc   3840

agctactcgg gaggctgagg tacgagaatc gcttgaacct gggaggcaga ggatgcagtg   3900

agccgagatc acgccattgc actccagcct gggggacaag agtgaatctg tgtctcacca   3960

aaaaaaaaaa gaaaaagaaa gatgcttaac aaaggttacc ataagccaca aattcataac   4020

cacttatcct tccagtttca agtagaatat attcataacc tcaataaagt tctccctgct   4080

cccaaa                                                              4086
```

<210> SEQ ID NO 57
<211> LENGTH: 339
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 57

```
Met Trp Gln Leu Trp Ala Ser Leu Cys Cys Leu Leu Val Leu Ala Asn
1               5                   10                  15

Ala Arg Ser Arg Pro Ser Phe His Pro Leu Ser Asp Glu Leu Val Asn
            20                  25                  30

Tyr Val Asn Lys Arg Asn Thr Thr Trp Gln Ala Gly His Asn Phe Tyr
        35                  40                  45

Asn Val Asp Met Ser Tyr Leu Lys Arg Leu Cys Gly Thr Phe Leu Gly
    50                  55                  60

Gly Pro Lys Pro Pro Gln Arg Val Met Phe Thr Glu Asp Leu Lys Leu
65                  70                  75                  80

Pro Ala Ser Phe Asp Ala Arg Glu Gln Trp Pro Gln Cys Pro Thr Ile
                85                  90                  95

Lys Glu Ile Arg Asp Gln Gly Ser Cys Gly Ser Cys Trp Ala Phe Gly
            100                 105                 110

Ala Val Glu Ala Ile Ser Asp Arg Ile Cys Ile His Thr Asn Ala His
        115                 120                 125

Val Ser Val Glu Val Ser Ala Glu Asp Leu Leu Thr Cys Cys Gly Ser
    130                 135                 140

Met Cys Gly Asp Gly Cys Asn Gly Gly Tyr Pro Ala Glu Ala Trp Asn
145                 150                 155                 160

Phe Trp Thr Arg Lys Gly Leu Val Ser Gly Gly Leu Tyr Glu Ser His
                165                 170                 175

Val Gly Cys Arg Pro Tyr Ser Ile Pro Pro Cys Glu His His Val Asn
            180                 185                 190

Gly Ser Arg Pro Pro Cys Thr Gly Glu Gly Asp Thr Pro Lys Cys Ser
        195                 200                 205

Lys Ile Cys Glu Pro Gly Tyr Ser Pro Thr Tyr Lys Gln Asp Lys His
    210                 215                 220

Tyr Gly Tyr Asn Ser Tyr Ser Val Ser Asn Ser Glu Lys Asp Ile Met
225                 230                 235                 240

Ala Glu Ile Tyr Lys Asn Gly Pro Val Glu Gly Ala Phe Ser Val Tyr
                245                 250                 255
```

-continued

```
Ser Asp Phe Leu Leu Tyr Lys Ser Gly Val Tyr Gln His Val Thr Gly
            260                 265                 270

Glu Met Met Gly Gly His Ala Ile Arg Ile Leu Gly Trp Gly Val Glu
        275                 280                 285

Asn Gly Thr Pro Tyr Trp Leu Val Ala Asn Ser Trp Asn Thr Asp Trp
    290                 295                 300

Gly Asp Asn Gly Phe Phe Lys Ile Leu Arg Gly Gln Asp His Cys Gly
305                 310                 315                 320

Ile Glu Ser Glu Val Val Ala Gly Ile Pro Arg Thr Asp Gln Tyr Trp
                325                 330                 335

Glu Lys Ile
```

<210> SEQ ID NO 58
<211> LENGTH: 1587
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 58

```
ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc     60

agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac    120

cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag    180

gacgcggtcg agtaggtttt aaaacatgaa tcctacactc atccttgctg ccttttgcct    240

gggaattgcc tcagctactc taacatttga tcacagttta gaggcacagt ggaccaagtg    300

gaaggcgatg cacaacagat tatacggcat gaatgaagaa ggatggagga gagcagtgtg    360

ggagaagaac atgaagatga ttgaactgca caatcaggaa tacagggaag ggaaacacag    420

cttcacaatg gccatgaacg cctttggaga catgaccagt gaagaattca ggcaggtgat    480

gaatggcttt caaaaccgta agcccaggaa ggggaaagtg ttccaggaac ctctgtttta    540

tgaggccccc agatctgtgg attggagaga gaaaggctac gtgactcctg tgaagaatca    600

gggtcagtgt ggttcttgtt gggcttttag tgctactggt gctcttgaag gacagatgtt    660

ccggaaaact gggaggctta tctcactgag tgagcagaat ctggtagact gctctgggcc    720

tcaaggcaat gaaggctgca atggtggcct aatggattat gctttccagt atgttcagga    780

taatggaggc ctggactctg aggaatccta tccatatgag gcaacagaag aatcctgtaa    840

gtacaatccc aagtattctg ttgctaatga caccggcttt gtggacatcc ctaagcagga    900

gaaggccctg atgaaggcag ttgcaactgt ggggcccatt tctgttgcta ttgatgcagg    960

tcatgagtcc ttcctgttct ataaagaagg catttatttt gagccagact gtagcagtga   1020

agacatggat catggtgtgc tggtggttgg ctacggattt gaaagcacag aatcagataa   1080

caataaatat tggctggtga agaacagctg gggtgaagaa tggggcatgg gtggctacgt   1140

aaagatggcc aaagaccgga gaaaccattg tggaattgcc tcagcagcca gctaccccac   1200

tgtgtgagct ggtggacggt gatgaggaag gacttgactg gggatggcgc atgcatggga   1260

ggaattcatc ttcagtctac cagcccccgc tgtgtcggat acacactcga atcattgaag   1320

atccgagtgt gatttgaatt ctgtgatatt ttcacactgg taaatgttac ctctatttta   1380

attactgcta taaataggtt tatattattg attcacttac tgactttgca ttttcgtttt   1440

taaaaggatg tataaatttt tacctgttta aataaaattt aatttcaaat gtagtggtgg   1500

ggcttctttc tatttttgat gcactgaatt tttgtgtaat aaagaacata attgggctct   1560
```

-continued

aagccataaa aaaaaaaaaa aaaaaaa 1587

<210> SEQ ID NO 59
<211> LENGTH: 333
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 59

Met Asn Pro Thr Leu Ile Leu Ala Ala Phe Cys Leu Gly Ile Ala Ser
1 5 10 15

Ala Thr Leu Thr Phe Asp His Ser Leu Glu Ala Gln Trp Thr Lys Trp
20 25 30

Lys Ala Met His Asn Arg Leu Tyr Gly Met Asn Glu Glu Gly Trp Arg
35 40 45

Arg Ala Val Trp Glu Lys Asn Met Lys Met Ile Glu Leu His Asn Gln
50 55 60

Glu Tyr Arg Glu Gly Lys His Ser Phe Thr Met Ala Met Asn Ala Phe
65 70 75 80

Gly Asp Met Thr Ser Glu Glu Phe Arg Gln Val Met Asn Gly Phe Gln
85 90 95

Asn Arg Lys Pro Arg Lys Gly Lys Val Phe Gln Glu Pro Leu Phe Tyr
100 105 110

Glu Ala Pro Arg Ser Val Asp Trp Arg Glu Lys Gly Tyr Val Thr Pro
115 120 125

Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ala Thr
130 135 140

Gly Ala Leu Glu Gly Gln Met Phe Arg Lys Thr Gly Arg Leu Ile Ser
145 150 155 160

Leu Ser Glu Gln Asn Leu Val Asp Cys Ser Gly Pro Gln Gly Asn Glu
165 170 175

Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Gln Tyr Val Gln Asp
180 185 190

Asn Gly Gly Leu Asp Ser Glu Glu Ser Tyr Pro Tyr Glu Ala Thr Glu
195 200 205

Glu Ser Cys Lys Tyr Asn Pro Lys Tyr Ser Val Ala Asn Asp Thr Gly
210 215 220

Phe Val Asp Ile Pro Lys Gln Glu Lys Ala Leu Met Lys Ala Val Ala
225 230 235 240

Thr Val Gly Pro Ile Ser Val Ala Ile Asp Ala Gly His Glu Ser Phe
245 250 255

Leu Phe Tyr Lys Glu Gly Ile Tyr Phe Glu Pro Asp Cys Ser Ser Glu
260 265 270

Asp Met Asp His Gly Val Leu Val Val Gly Tyr Gly Phe Glu Ser Thr
275 280 285

Glu Ser Asp Asn Asn Lys Tyr Trp Leu Val Lys Asn Ser Trp Gly Glu
290 295 300

Glu Trp Gly Met Gly Gly Tyr Val Lys Met Ala Lys Asp Arg Arg Asn
305 310 315 320

His Cys Gly Ile Ala Ser Ala Ala Ser Tyr Pro Thr Val
325 330

<210> SEQ ID NO 60

-continued

<211> LENGTH: 1626
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 60

ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc 60

agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac 120

cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag 180

gacgcggtcg agtaggtgtg caccagccct ggcaacgaga gcgtctaccc cgaactctgc 240

tggccttgag gttttaaaac atgaatccta cactcatcct tgctgccttt tgcctgggaa 300

ttgcctcagc tactctaaca tttgatcaca gtttagaggc acagtggacc aagtggaagg 360

cgatgcacaa cagattatac ggcatgaatg aagaaggatg gaggagagca gtgtgggaga 420

agaacatgaa gatgattgaa ctgcacaatc aggaatacag ggaagggaaa cacagcttca 480

caatggccat gaacgccttt ggagacatga ccagtgaaga attcaggcag gtgatgaatg 540

gctttcaaaa ccgtaagccc aggaagggga aagtgttcca ggaacctctg ttttatgagg 600

cccccagatc tgtggattgg agagagaaag gctacgtgac tcctgtgaag aatcagggtc 660

agtgtggttc ttgttgggct tttagtgcta ctggtgctct tgaaggacag atgttccgga 720

aaactgggag gcttatctca ctgagtgagc agaatctggt agactgctct gggcctcaag 780

gcaatgaagg ctgcaatggt ggcctaatgg attatgcttt ccagtatgtt caggataatg 840

gaggcctgga ctctgaggaa tcctatccat atgaggcaac agaagaatcc tgtaagtaca 900

atcccaagta ttctgttgct aatgacaccg gctttgtgga catccctaag caggagaagg 960

ccctgatgaa ggcagttgca actgtggggc ccatttctgt tgctattgat gcaggtcatg 1020

agtccttcct gttctataaa gaaggcattt attttgagcc agactgtagc agtgaagaca 1080

tggatcatgg tgtgctggtg gttggctacg gatttgaaag cacagaatca gataacaata 1140

aatattggct ggtgaagaac agctggggtg aagaatgggg catgggtggc tacgtaaaga 1200

tggccaaaga ccggagaaac cattgtggaa ttgcctcagc agccagctac cccactgtgt 1260

gagctggtgg acggtgatga ggaaggactt gactggggat ggcgcatgca tgggaggaat 1320

tcatcttcag tctaccagcc cccgctgtgt cggatacaca ctcgaatcat tgaagatccg 1380

agtgtgattt gaattctgtg atattttcac actggtaaat gttacctcta ttttaattac 1440

tgctataaat aggtttatat tattgattca cttactgact ttgcattttc gtttttaaaa 1500

ggatgtataa atttttacct gtttaaataa aatttaattt caaatgtagt ggtggggctt 1560

ctttctattt ttgatgcact gaatttttgt gtaataaaga acataattgg gctctaagcc 1620

ataaaa 1626

<210> SEQ ID NO 61
<211> LENGTH: 333
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 61

Met Asn Pro Thr Leu Ile Leu Ala Ala Phe Cys Leu Gly Ile Ala Ser
1 5 10 15

Ala Thr Leu Thr Phe Asp His Ser Leu Glu Ala Gln Trp Thr Lys Trp
20 25 30

-continued

```
Lys Ala Met His Asn Arg Leu Tyr Gly Met Asn Glu Glu Gly Trp Arg
        35                  40                  45

Arg Ala Val Trp Glu Lys Asn Met Lys Met Ile Glu Leu His Asn Gln
    50                  55                  60

Glu Tyr Arg Glu Gly Lys His Ser Phe Thr Met Ala Met Asn Ala Phe
65                  70                  75                  80

Gly Asp Met Thr Ser Glu Glu Phe Arg Gln Val Met Asn Gly Phe Gln
                85                  90                  95

Asn Arg Lys Pro Arg Lys Gly Lys Val Phe Gln Glu Pro Leu Phe Tyr
            100                 105                 110

Glu Ala Pro Arg Ser Val Asp Trp Arg Glu Lys Gly Tyr Val Thr Pro
        115                 120                 125

Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ala Thr
    130                 135                 140

Gly Ala Leu Glu Gly Gln Met Phe Arg Lys Thr Gly Arg Leu Ile Ser
145                 150                 155                 160

Leu Ser Glu Gln Asn Leu Val Asp Cys Ser Gly Pro Gln Gly Asn Glu
                165                 170                 175

Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Gln Tyr Val Gln Asp
            180                 185                 190

Asn Gly Gly Leu Asp Ser Glu Glu Ser Tyr Pro Tyr Glu Ala Thr Glu
        195                 200                 205

Glu Ser Cys Lys Tyr Asn Pro Lys Tyr Ser Val Ala Asn Asp Thr Gly
    210                 215                 220

Phe Val Asp Ile Pro Lys Gln Glu Lys Ala Leu Met Lys Ala Val Ala
225                 230                 235                 240

Thr Val Gly Pro Ile Ser Val Ala Ile Asp Ala Gly His Glu Ser Phe
                245                 250                 255

Leu Phe Tyr Lys Glu Gly Ile Tyr Phe Glu Pro Asp Cys Ser Ser Glu
            260                 265                 270

Asp Met Asp His Gly Val Leu Val Val Gly Tyr Gly Phe Glu Ser Thr
        275                 280                 285

Glu Ser Asp Asn Asn Lys Tyr Trp Leu Val Lys Asn Ser Trp Gly Glu
    290                 295                 300

Glu Trp Gly Met Gly Gly Tyr Val Lys Met Ala Lys Asp Arg Arg Asn
305                 310                 315                 320

His Cys Gly Ile Ala Ser Ala Ala Ser Tyr Pro Thr Val
                325                 330
```

<210> SEQ ID NO 62
<211> LENGTH: 1567
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 62

```
ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc     60

agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac    120

cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag    180

gacgcggtcg agttttaaaa catgaatcct acactcatcc ttgctgcctt ttgcctggga    240

attgcctcag ctactctaac atttgatcac agtttagagg cacagtggac caagtggaag    300

gcgatgcaca acagattata cggcatgaat gaagaaggat ggaggagagc agtgtgggag    360
```

-continued

aagaacatga agatgattga actgcacaat caggaataca gggaagggaa acacagcttc 420

acaatggcca tgaacgcctt tggagacatg accagtgaag aattcaggca ggtgatgaat 480

ggctttcaaa accgtaagcc caggaagggg aaagtgttcc aggaacctct gttttatgag 540

gcccccagat ctgtggattg gagagagaaa ggctacgtga ctcctgtgaa gaatcagggt 600

cagtgtggtt cttgttgggc ttttagtgct actggtgctc ttgaaggaca gatgttccgg 660

aaaactggga ggcttatctc actgagtgag cagaatctgg tagactgctc tgggcctcaa 720

ggcaatgaag gctgcaatgg tggcctaatg gattatgctt tccagtatgt tcaggataat 780

ggaggcctgg actctgagga atcctatcca tatgaggcaa cagaagaatc ctgtaagtac 840

aatcccaagt attctgttgc taatgacacc ggctttgtgg acatccctaa gcaggagaag 900

gccctgatga aggcagttgc aactgtgggg cccatttctg ttgctattga tgcaggtcat 960

gagtccttcc tgttctataa agaaggcatt tattttgagc cagactgtag cagtgaagac 1020

atggatcatg gtgtgctggt ggttggctac ggatttgaaa gcacagaatc agataacaat 1080

aaatattggc tggtgaagaa cagctggggt gaagaatggg gcatgggtgg ctacgtaaag 1140

atggccaaag accggagaaa ccattgtgga attgcctcag cagccagcta ccccactgtg 1200

tgagctggtg gacggtgatg aggaaggact tgactgggga tggcgcatgc atgggaggaa 1260

ttcatcttca gtctaccagc ccccgctgtg tcggatacac actcgaatca ttgaagatcc 1320

gagtgtgatt tgaattctgt gatattttca cactggtaaa tgttacctct attttaatta 1380

ctgctataaa taggtttata ttattgattc acttactgac tttgcatttt cgtttttaaa 1440

aggatgtata aatttttacc tgtttaaata aaatttaatt tcaaatgtag tggtggggct 1500

tctttctatt tttgatgcac tgaatttttg tgtaataaag aacataattg ggctctaagc 1560

cataaaa 1567

<210> SEQ ID NO 63
<211> LENGTH: 333
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 63

Met Asn Pro Thr Leu Ile Leu Ala Ala Phe Cys Leu Gly Ile Ala Ser
1 5 10 15

Ala Thr Leu Thr Phe Asp His Ser Leu Glu Ala Gln Trp Thr Lys Trp
20 25 30

Lys Ala Met His Asn Arg Leu Tyr Gly Met Asn Glu Glu Gly Trp Arg
35 40 45

Arg Ala Val Trp Glu Lys Asn Met Lys Met Ile Glu Leu His Asn Gln
50 55 60

Glu Tyr Arg Glu Gly Lys His Ser Phe Thr Met Ala Met Asn Ala Phe
65 70 75 80

Gly Asp Met Thr Ser Glu Glu Phe Arg Gln Val Met Asn Gly Phe Gln
85 90 95

Asn Arg Lys Pro Arg Lys Gly Lys Val Phe Gln Glu Pro Leu Phe Tyr
100 105 110

Glu Ala Pro Arg Ser Val Asp Trp Arg Glu Lys Gly Tyr Val Thr Pro
115 120 125

Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ala Thr

-continued

```
    130                 135                 140

Gly Ala Leu Glu Gly Gln Met Phe Arg Lys Thr Gly Arg Leu Ile Ser
145                 150                 155                 160

Leu Ser Glu Gln Asn Leu Val Asp Cys Ser Gly Pro Gln Gly Asn Glu
                165                 170                 175

Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Gln Tyr Val Gln Asp
            180                 185                 190

Asn Gly Gly Leu Asp Ser Glu Glu Ser Tyr Pro Tyr Glu Ala Thr Glu
        195                 200                 205

Glu Ser Cys Lys Tyr Asn Pro Lys Tyr Ser Val Ala Asn Asp Thr Gly
    210                 215                 220

Phe Val Asp Ile Pro Lys Gln Glu Lys Ala Leu Met Lys Ala Val Ala
225                 230                 235                 240

Thr Val Gly Pro Ile Ser Val Ala Ile Asp Ala Gly His Glu Ser Phe
                245                 250                 255

Leu Phe Tyr Lys Glu Gly Ile Tyr Phe Glu Pro Asp Cys Ser Ser Glu
            260                 265                 270

Asp Met Asp His Gly Val Leu Val Val Gly Tyr Gly Phe Glu Ser Thr
        275                 280                 285

Glu Ser Asp Asn Asn Lys Tyr Trp Leu Val Lys Asn Ser Trp Gly Glu
    290                 295                 300

Glu Trp Gly Met Gly Gly Tyr Val Lys Met Ala Lys Asp Arg Arg Asn
305                 310                 315                 320

His Cys Gly Ile Ala Ser Ala Ala Ser Tyr Pro Thr Val
                325                 330
```

<210> SEQ ID NO 64
<211> LENGTH: 1141
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 64

```
ggcggtgccg gccgaaccca gacccgaggt tttagaagca gagtcaggcg aagctgggcc     60

agaaccgcga cctccgcaac cttgagcggc atccgtggag tgcgcctgcg cagctacgac    120

cgcagcagga aagcgccgcc ggccaggccc agctgtggcc ggacagggac tggaagagag    180

gacgcggtcg agtaggtttt aaaacatgaa tcctacactc atccttgctg ccttttgcct    240

gggaattgcc tcagctactc taacatttga tcacagttta gaggcacagt ggaccaagtg    300

gaaggctgca atggtggcct aatggattat gctttccagt atgttcagga taatggaggc    360

ctggactctg aggaatccta tccatatgag gcaacagaag aatcctgtaa gtacaatccc    420

aagtattctg ttgctaatga caccggcttt gtggacatcc ctaagcagga gaaggccctg    480

atgaaggcag ttgcaactgt ggggcccatt tctgttgcta ttgatgcagg tcatgagtcc    540

ttcctgttct ataaagaagg catttatttt gagccagact gtagcagtga agacatggat    600

catggtgtgc tggtggttgg ctacggattt gaaagcacag aatcagataa caataaatat    660

tggctggtga agaacagctg gggtgaagaa tggggcatgg gtggctacgt aaagatggcc    720

aaagaccgga gaaaccattg tggaattgcc tcagcagcca gctaccccac tgtgtgagct    780

ggtggacggt gatgaggaag gacttgactg gggatggcgc atgcatggga ggaattcatc    840

ttcagtctac cagcccccgc tgtgtcggat acacactcga atcattgaag atccgagtgt    900
```

-continued

gatttgaatt ctgtgatatt ttcacactgg taaatgttac ctctatttta attactgcta 960

taaataggtt tatattattg attcacttac tgactttgca ttttcgtttt taaaaggatg 1020

tataaatttt tacctgttta aataaaattt aatttcaaat gtagtggtgg ggcttctttc 1080

tatttttgat gcactgaatt tttgtgtaat aaagaacata attgggctct aagccataaa 1140

a 1141

<210> SEQ ID NO 65
<211> LENGTH: 151
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 65

Met Asp Tyr Ala Phe Gln Tyr Val Gln Asp Asn Gly Gly Leu Asp Ser
1 5 10 15

Glu Glu Ser Tyr Pro Tyr Glu Ala Thr Glu Glu Ser Cys Lys Tyr Asn
20 25 30

Pro Lys Tyr Ser Val Ala Asn Asp Thr Gly Phe Val Asp Ile Pro Lys
35 40 45

Gln Glu Lys Ala Leu Met Lys Ala Val Ala Thr Val Gly Pro Ile Ser
50 55 60

Val Ala Ile Asp Ala Gly His Glu Ser Phe Leu Phe Tyr Lys Glu Gly
65 70 75 80

Ile Tyr Phe Glu Pro Asp Cys Ser Ser Glu Asp Met Asp His Gly Val
85 90 95

Leu Val Val Gly Tyr Gly Phe Glu Ser Thr Glu Ser Asp Asn Asn Lys
100 105 110

Tyr Trp Leu Val Lys Asn Ser Trp Gly Glu Glu Trp Gly Met Gly Gly
115 120 125

Tyr Val Lys Met Ala Lys Asp Arg Arg Asn His Cys Gly Ile Ala Ser
130 135 140

Ala Ala Ser Tyr Pro Thr Val
145 150

<210> SEQ ID NO 66
<211> LENGTH: 1401
<212> TYPE: DNA
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 66

acagctctgg acaggctgct tttcattttg gtgagtccat ccagtacctc cacgtgccct 60

gtttttctcc aggcacatcc ttggcctctt ccacagtcct tgggttttaa aacatgaatc 120

ctacactcat ccttgctgcc ttttgcctgg gaattgcctc agctactcta acatttgatc 180

acagtttaga ggcacagtgg accaagtgga aggcgatgca caacagatta tacggcatga 240

atgaagaagg atggaggaga gcagtgtggg agaagaacat gaagatgatt gaactgcaca 300

atcaggaata cagggaaggg aaacacagct tcacaatggc catgaacgcc tttggagaca 360

tgaccagtga agaattcagg caggtgatga atggctttca aaaccgtaag cccaggaagg 420

ggaaagtgtt ccaggaacct ctgttttatg aggcccccag atctgtggat tggagagaga 480

aaggctacgt gactcctgtg aagaatcagg gtcagtgtgg ttcttgttgg gcttttagtg 540

ctactggtgc tcttgaagga cagatgttcc ggaaaactgg gaggcttatc tcactgagtg 600

-continued

```
agcagaatct ggtagactgc tctgggcctc aaggcaatga aggctgcaat ggtggcctaa    660

tggattatgc tttccagtat gttcaggata atggaggcct ggactctgag gaatcctatc    720

catatgaggc aacagaagaa tcctgtaagt acaatcccaa gtattctgtt gctaatgaca    780

ccggctttgt ggacatccct aagcaggaga aggccctgat gaaggcagtt gcaactgtgg    840

ggcccatttc tgttgctatt gatgcaggtc atgagtcctt cctgttctat aaagaaggca    900

tttattttga gccagactgt agcagtgaag acatggatca tggtgtgctg gtggttggct    960

acggatttga aagcacagaa tcagataaca ataaatattg gctggtgaag aacagctggg   1020

gtgaagaatg gggcatgggt ggctacgtaa agatggccaa agaccggaga aaccattgtg   1080

gaattgcctc agcagccagc taccccactg tgtgagctgg tggacggtga tgaggaagga   1140

cttgactggg gatggcgcat gcatgggagg aattcatctt cagtctacca gcccccgctg   1200

tgtcggatac acactcgaat cattgaagat ccgagtgtga tttgaattct gtgatatttt   1260

cacactggta aatgttacct ctattttaat tactgctata aataggttta tattattgat   1320

tcacttactg actttgcatt ttcgttttta aaaggatgta taaattttta cctgtttaaa   1380

taaaatttaa tttcaaatgt a                                             1401
```

<210> SEQ ID NO 67
<211> LENGTH: 333
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 67

```
Met Asn Pro Thr Leu Ile Leu Ala Ala Phe Cys Leu Gly Ile Ala Ser
1               5                  10                  15

Ala Thr Leu Thr Phe Asp His Ser Leu Glu Ala Gln Trp Thr Lys Trp
            20                  25                  30

Lys Ala Met His Asn Arg Leu Tyr Gly Met Asn Glu Glu Gly Trp Arg
        35                  40                  45

Arg Ala Val Trp Glu Lys Asn Met Lys Met Ile Glu Leu His Asn Gln
    50                  55                  60

Glu Tyr Arg Glu Gly Lys His Ser Phe Thr Met Ala Met Asn Ala Phe
65                  70                  75                  80

Gly Asp Met Thr Ser Glu Glu Phe Arg Gln Val Met Asn Gly Phe Gln
                85                  90                  95

Asn Arg Lys Pro Arg Lys Gly Lys Val Phe Gln Glu Pro Leu Phe Tyr
            100                 105                 110

Glu Ala Pro Arg Ser Val Asp Trp Arg Glu Lys Gly Tyr Val Thr Pro
        115                 120                 125

Val Lys Asn Gln Gly Gln Cys Gly Ser Cys Trp Ala Phe Ser Ala Thr
    130                 135                 140

Gly Ala Leu Glu Gly Gln Met Phe Arg Lys Thr Gly Arg Leu Ile Ser
145                 150                 155                 160

Leu Ser Glu Gln Asn Leu Val Asp Cys Ser Gly Pro Gln Gly Asn Glu
                165                 170                 175

Gly Cys Asn Gly Gly Leu Met Asp Tyr Ala Phe Gln Tyr Val Gln Asp
            180                 185                 190

Asn Gly Gly Leu Asp Ser Glu Glu Ser Tyr Pro Tyr Glu Ala Thr Glu
        195                 200                 205

Glu Ser Cys Lys Tyr Asn Pro Lys Tyr Ser Val Ala Asn Asp Thr Gly
```

-continued

```
    210                 215                 220

Phe Val Asp Ile Pro Lys Gln Glu Lys Ala Leu Met Lys Ala Val Ala
225                 230                 235                 240

Thr Val Gly Pro Ile Ser Val Ala Ile Asp Ala Gly His Glu Ser Phe
                245                 250                 255

Leu Phe Tyr Lys Glu Gly Ile Tyr Phe Glu Pro Asp Cys Ser Ser Glu
            260                 265                 270

Asp Met Asp His Gly Val Leu Val Val Gly Tyr Gly Phe Glu Ser Thr
        275                 280                 285

Glu Ser Asp Asn Asn Lys Tyr Trp Leu Val Lys Asn Ser Trp Gly Glu
    290                 295                 300

Glu Trp Gly Met Gly Gly Tyr Val Lys Met Ala Lys Asp Arg Arg Asn
305                 310                 315                 320

His Cys Gly Ile Ala Ser Ala Ala Ser Tyr Pro Thr Val
                325                 330


<210> SEQ ID NO 68
<211> LENGTH: 412
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 68

Met Gln Pro Ser Ser Leu Leu Pro Leu Ala Leu Cys Leu Leu Ala Ala
1               5                   10                  15

Pro Ala Ser Ala Leu Val Arg Ile Pro Leu His Lys Phe Thr Ser Ile
            20                  25                  30

Arg Arg Thr Met Ser Glu Val Gly Gly Ser Val Glu Asp Leu Ile Ala
        35                  40                  45

Lys Gly Pro Val Ser Lys Tyr Ser Gln Ala Val Pro Ala Val Thr Glu
    50                  55                  60

Gly Pro Ile Pro Glu Val Leu Lys Asn Tyr Met Asp Ala Gln Tyr Tyr
65                  70                  75                  80

Gly Glu Ile Gly Ile Gly Thr Pro Pro Gln Cys Phe Thr Val Val Phe
                85                  90                  95

Asp Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Ile His Cys Lys Leu
            100                 105                 110

Leu Asp Ile Ala Cys Trp Ile His His Lys Tyr Asn Ser Asp Lys Ser
        115                 120                 125

Ser Thr Tyr Val Lys Asn Gly Thr Ser Phe Asp Ile His Tyr Gly Ser
    130                 135                 140

Gly Ser Leu Ser Gly Tyr Leu Ser Gln Asp Thr Val Ser Val Pro Cys
145                 150                 155                 160

Gln Ser Ala Ser Ser Ala Ser Ala Leu Gly Gly Val Lys Val Glu Arg
                165                 170                 175

Gln Val Phe Gly Glu Ala Thr Lys Gln Pro Gly Ile Thr Phe Ile Ala
            180                 185                 190

Ala Lys Phe Asp Gly Ile Leu Gly Met Ala Tyr Pro Arg Ile Ser Val
        195                 200                 205

Asn Asn Val Leu Pro Val Phe Asp Asn Leu Met Gln Gln Lys Leu Val
    210                 215                 220

Asp Gln Asn Ile Phe Ser Phe Tyr Leu Ser Arg Asp Pro Asp Ala Gln
225                 230                 235                 240
```

-continued

```
Pro Gly Gly Glu Leu Met Leu Gly Gly Thr Asp Ser Lys Tyr Tyr Lys
                245                 250                 255

Gly Ser Leu Ser Tyr Leu Asn Val Thr Arg Lys Ala Tyr Trp Gln Val
            260                 265                 270

His Leu Asp Gln Val Glu Val Ala Ser Gly Leu Thr Leu Cys Lys Glu
        275                 280                 285

Gly Cys Glu Ala Ile Val Asp Thr Gly Thr Ser Leu Met Val Gly Pro
    290                 295                 300

Val Asp Glu Val Arg Glu Leu Gln Lys Ala Ile Gly Ala Val Pro Leu
305                 310                 315                 320

Ile Gln Gly Glu Tyr Met Ile Pro Cys Glu Lys Val Ser Thr Leu Pro
                325                 330                 335

Ala Ile Thr Leu Lys Leu Gly Gly Lys Gly Tyr Lys Leu Ser Pro Glu
            340                 345                 350

Asp Tyr Thr Leu Lys Val Ser Gln Ala Gly Lys Thr Leu Cys Leu Ser
        355                 360                 365

Gly Phe Met Gly Met Asp Ile Pro Pro Pro Ser Gly Pro Leu Trp Ile
    370                 375                 380

Leu Gly Asp Val Phe Ile Gly Arg Tyr Tyr Thr Val Phe Asp Arg Asp
385                 390                 395                 400

Asn Asn Arg Val Gly Phe Ala Glu Ala Ala Arg Leu
                405                 410


<210> SEQ ID NO 69
<211> LENGTH: 401
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 69

Met Lys Thr Leu Leu Leu Leu Leu Leu Val Leu Leu Glu Leu Gly Glu
1               5                   10                  15

Ala Gln Gly Ser Leu His Arg Val Pro Leu Arg Arg His Pro Ser Leu
            20                  25                  30

Lys Lys Lys Leu Arg Ala Arg Ser Gln Leu Ser Glu Phe Trp Lys Ser
        35                  40                  45

His Asn Leu Asp Met Ile Gln Phe Thr Glu Ser Cys Ser Met Asp Gln
    50                  55                  60

Ser Ala Lys Glu Pro Leu Ile Asn Tyr Leu Asp Met Glu Tyr Phe Gly
65                  70                  75                  80

Thr Ile Ser Ile Gly Ser Pro Pro Gln Asn Phe Thr Val Ile Phe Asp
                85                  90                  95

Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Val Tyr Cys Thr Ser Pro
            100                 105                 110

Ala Cys Lys Thr His Ser Arg Phe Gln Pro Ser Gln Ser Ser Thr Tyr
        115                 120                 125

Ser Gln Pro Gly Gln Ser Phe Ser Ile Gln Tyr Gly Thr Gly Ser Leu
    130                 135                 140

Ser Gly Ile Ile Gly Ala Asp Gln Val Ser Ala Phe Ala Thr Gln Val
145                 150                 155                 160

Glu Gly Leu Thr Val Val Gly Gln Gln Phe Gly Glu Ser Val Thr Glu
                165                 170                 175

Pro Gly Gln Thr Phe Val Asp Ala Glu Phe Asp Gly Ile Leu Gly Leu
            180                 185                 190
```

-continued

```
Gly Tyr Pro Ser Leu Ala Val Gly Gly Val Thr Pro Val Phe Asp Asn
        195                 200                 205

Met Met Ala Gln Asn Leu Val Asp Leu Pro Met Phe Ser Val Tyr Met
    210                 215                 220

Ser Ser Asn Pro Glu Gly Gly Ala Gly Ser Glu Leu Ile Phe Gly Gly
225                 230                 235                 240

Tyr Asp His Ser His Phe Ser Gly Ser Leu Asn Trp Val Pro Val Thr
                245                 250                 255

Lys Gln Ala Tyr Trp Gln Ile Ala Leu Asp Asn Ile Gln Val Gly Gly
            260                 265                 270

Thr Val Met Phe Cys Ser Glu Gly Cys Gln Ala Ile Val Asp Thr Gly
        275                 280                 285

Thr Ser Leu Ile Thr Gly Pro Ser Asp Lys Ile Lys Gln Leu Gln Asn
    290                 295                 300

Ala Ile Gly Ala Ala Pro Val Asp Gly Glu Tyr Ala Val Glu Cys Ala
305                 310                 315                 320

Asn Leu Asn Val Met Pro Asp Val Thr Phe Thr Ile Asn Gly Val Pro
                325                 330                 335

Tyr Thr Leu Ser Pro Thr Ala Tyr Thr Leu Leu Asp Phe Val Asp Gly
            340                 345                 350

Met Gln Phe Cys Ser Ser Gly Phe Gln Gly Leu Asp Ile His Pro Pro
        355                 360                 365

Ala Gly Pro Leu Trp Ile Leu Gly Asp Val Phe Ile Arg Gln Phe Tyr
    370                 375                 380

Ser Val Phe Asp Arg Gly Asn Asn Arg Val Gly Leu Ala Pro Ala Val
385                 390                 395                 400

Pro
```

<210> SEQ ID NO 70
<211> LENGTH: 396
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 70

```
Met Lys Thr Leu Leu Leu Leu Leu Leu Val Leu Leu Glu Leu Gly Glu
1               5                   10                  15

Ala Gln Gly Ser Leu His Arg Val Pro Leu Arg Arg His Pro Ser Leu
            20                  25                  30

Lys Lys Lys Leu Arg Ala Arg Ser Gln Leu Ser Glu Phe Trp Lys Ser
        35                  40                  45

His Asn Leu Asp Met Ile Gln Phe Thr Glu Ser Cys Ser Met Asp Gln
    50                  55                  60

Ser Ala Lys Glu Pro Leu Ile Asn Tyr Leu Asp Met Glu Tyr Phe Gly
65                  70                  75                  80

Thr Ile Ser Ile Gly Ser Pro Pro Gln Asn Phe Thr Val Ile Phe Asp
                85                  90                  95

Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Val Tyr Cys Thr Ser Pro
            100                 105                 110

Ala Cys Lys Thr His Ser Arg Phe Gln Pro Ser Gln Ser Ser Thr Tyr
        115                 120                 125

Ser Gln Pro Gly Gln Ser Phe Ser Ile Gln Tyr Gly Thr Gly Ser Leu
    130                 135                 140
```

-continued

```
Ser Gly Ile Ile Gly Ala Asp Gln Val Ser Val Glu Gly Leu Thr Val
145                 150                 155                 160

Val Gly Gln Gln Phe Gly Glu Ser Val Thr Glu Pro Gly Gln Thr Phe
                165                 170                 175

Val Asp Ala Glu Phe Asp Gly Ile Leu Gly Leu Gly Tyr Pro Ser Leu
            180                 185                 190

Ala Val Gly Gly Val Thr Pro Val Phe Asp Asn Met Met Ala Gln Asn
        195                 200                 205

Leu Val Asp Leu Pro Met Phe Ser Val Tyr Met Ser Ser Asn Pro Glu
    210                 215                 220

Gly Gly Ala Gly Ser Glu Leu Ile Phe Gly Gly Tyr Asp His Ser His
225                 230                 235                 240

Phe Ser Gly Ser Leu Asn Trp Val Pro Val Thr Lys Gln Ala Tyr Trp
                245                 250                 255

Gln Ile Ala Leu Asp Asn Ile Gln Val Gly Gly Thr Val Met Phe Cys
            260                 265                 270

Ser Glu Gly Cys Gln Ala Ile Val Asp Thr Gly Thr Ser Leu Ile Thr
        275                 280                 285

Gly Pro Ser Asp Lys Ile Lys Gln Leu Gln Asn Ala Ile Gly Ala Ala
    290                 295                 300

Pro Val Asp Gly Glu Tyr Ala Val Glu Cys Ala Asn Leu Asn Val Met
305                 310                 315                 320

Pro Asp Val Thr Phe Thr Ile Asn Gly Val Pro Tyr Thr Leu Ser Pro
                325                 330                 335

Thr Ala Tyr Thr Leu Leu Asp Phe Val Asp Gly Met Gln Phe Cys Ser
            340                 345                 350

Ser Gly Phe Gln Gly Leu Asp Ile His Pro Pro Ala Gly Pro Leu Trp
        355                 360                 365

Ile Leu Gly Asp Val Phe Ile Arg Gln Phe Tyr Ser Val Phe Asp Arg
    370                 375                 380

Gly Asn Asn Arg Val Gly Leu Ala Pro Ala Val Pro
385                 390                 395


<210> SEQ ID NO 71
<211> LENGTH: 363
<212> TYPE: PRT
<213> ORGANISM: Homo sapiens

<400> SEQUENCE: 71

Met Lys Thr Leu Leu Leu Leu Leu Leu Val Leu Leu Glu Leu Gly Glu
1               5                   10                  15

Ala Gln Gly Ser Leu His Arg Val Pro Leu Arg Arg His Pro Ser Leu
            20                  25                  30

Lys Lys Lys Leu Arg Ala Arg Ser Gln Leu Ser Glu Phe Trp Lys Ser
        35                  40                  45

His Asn Leu Asp Met Ile Gln Phe Thr Glu Ser Cys Ser Met Asp Gln
    50                  55                  60

Ser Ala Lys Glu Pro Leu Ile Asn Tyr Leu Asp Met Glu Tyr Phe Gly
65                  70                  75                  80

Thr Ile Ser Ile Gly Ser Pro Pro Gln Asn Phe Thr Val Ile Phe Asp
                85                  90                  95

Thr Gly Ser Ser Asn Leu Trp Val Pro Ser Val Tyr Cys Thr Ser Pro
```

-continued

```
            100                 105                 110

Ala Cys Lys Thr His Ser Arg Phe Gln Pro Ser Gln Ser Ser Thr Tyr
        115                 120                 125

Ser Gln Pro Gly Gln Ser Phe Ser Ile Gln Tyr Gly Thr Gly Ser Leu
    130                 135                 140

Ser Gly Ile Ile Gly Ala Asp Gln Val Ser Val Glu Gly Leu Thr Val
145                 150                 155                 160

Val Gly Gln Gln Phe Gly Glu Ser Val Thr Glu Pro Gly Gln Thr Phe
                165                 170                 175

Val Asp Ala Glu Phe Asp Gly Ile Leu Gly Leu Gly Tyr Pro Ser Leu
            180                 185                 190

Ala Val Gly Gly Val Thr Pro Val Phe Asp Asn Met Met Ala Gln Asn
        195                 200                 205

Leu Val Asp Leu Pro Met Phe Ser Val Tyr Met Ser Ser Asn Pro Glu
    210                 215                 220

Gly Gly Ala Gly Ser Glu Leu Ile Phe Gly Gly Tyr Asp His Ser His
225                 230                 235                 240

Phe Ser Gly Ser Leu Asn Trp Val Pro Val Thr Lys Gln Ala Tyr Trp
                245                 250                 255

Gln Ile Ala Leu Asp Asn Met Leu Trp Ser Val Pro Thr Leu Thr Ser
            260                 265                 270

Cys Arg Met Ser Pro Ser Pro Leu Thr Glu Ser Pro Ile Pro Ser Ala
        275                 280                 285

Gln Leu Pro Thr Pro Tyr Trp Thr Ser Trp Met Glu Cys Ser Ser Ala
    290                 295                 300

Ala Val Ala Phe Lys Asp Leu Thr Ser Thr Leu Gln Leu Gly Pro Ser
305                 310                 315                 320

Gly Ser Trp Gly Met Ser Ser Phe Asp Ser Phe Thr Gln Ser Leu Thr
                325                 330                 335

Val Gly Ile Thr Val Trp Asp Trp Pro Gln Gln Ser Pro Lys Glu Gly
            340                 345                 350

Pro Cys Val Cys Ala Cys Leu Ser Asp Arg Pro
        355                 360


<210> SEQ ID NO 72
<211> LENGTH: 12
<212> TYPE: PRT
<213> ORGANISM: Artificial Sequence
<220> FEATURE:
<223> OTHER INFORMATION: Substrate competitive inhibitor, L803-mts
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (11)..(11)
<223> OTHER INFORMATION: May be N-terminally myristoylated
<220> FEATURE:
<221> NAME/KEY: MISC_FEATURE
<222> LOCATION: (11)..(11)
<223> OTHER INFORMATION: May be a phosphorylated residue

<400> SEQUENCE: 72

Gly Lys Glu Ala Pro Pro Ala Pro Pro Gln Ser Pro
1               5                   10
```

1. A method of treating or preventing amyloidosis in a subject comprising administering to the subject a composition comprising a therapeutically effective amount of at least one catabolic enzyme or a biologically active fragment thereof.

2. The method of claim **1**, wherein the catabolic enzyme is selected from protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L.

3. The method of claim **2**, wherein the catabolic enzyme is PPCA, or a biologically active fragment thereof.

4. The method of claim **3**, wherein the PPCA polypeptide comprises an amino acid sequence with at least 85% sequence identity to SEQ ID NO: 2, 43, or 45, or a biologically active fragment thereof.

5. The method of claim **4**, wherein administration of the PPCA polypeptide comprises administration of a viral vector comprising a nucleotide sequence having at least 85% identity to SEQ ID NO: 1, 42, or 44.

6.-**13**. (canceled)

14. The method of claim **1**, wherein at least two catabolic enzymes are administered.

15. The method of claim **14**, wherein the catabolic enzymes are selected from protective protein/cathepsin A (PPCA), neuraminidase 1 (NEU1), tripeptidyl peptidase 1 (TPP1), cathepsin B, cathepsin D, cathepsin E, cathepsin K, and cathepsin L.

16. The method of claim **15**, wherein the catabolic enzymes are PPCA and NEU1.

17. (canceled)

18. The method of claim **1**, wherein the catabolic enzyme acts to prevent the formation of and/or degrade amyloid within the lysosome.

19. The method of claim **1**, wherein the catabolic enzyme is targeted to the cell lysosome.

20. The method of claim **1**, wherein the catabolic enzyme acts to prevent the accumulation of and/or degrade amyloid outside the cell.

21.-**24**. (canceled)

25. The method of claim **1**, wherein the subject is a human.

26-**27**. (canceled)

28. The method of claim **1**, wherein the amyloidosis is light-chain (AL) amyloidosis.

29. The method of claim **28**, wherein the AL amyloidosis involves one or more organs selected from the heart, the kidneys, the nervous system, and the gastrointestinal tract.

30. The method of claim **1**, wherein the amyloidosis is amyloid-beta (Aβ) amyloidosis.

31. The method of claim **30**, wherein the Aβ amyloidosis is associated one or more diseases selected from Alzheimer's disease, cerebral amyloid angiopathy, Lewy body dementia, and inclusion body myositis.

32. The method of claim **1**, further comprising the administration of one or more additional drugs for treating or preventing amyloidosis.

33. The method of claim **32**, wherein the one or more additional drugs is selected from melphalan, dexamethasone, prednisone, bortezomib, lenalidomide, vincristine, doxorubicin, and cyclophosphamide.

34. The method of claim **1**, further comprising the administration of one or more drugs that acidifies the lysosome.

35. The method of claim **34**, wherein the drug that acidifies the lysosome is selected from an acidic nanoparticle, a catecholamine, a β-adrenergic receptor agonist, an adenosine receptor agonist, a dopamine receptor agonist, an activator of the cystic fibrosis transmembrane conductance regulator (CFTR), cyclic adenosine monophosphate (cAMP), a cAMP analog, and an inhibitor of glycogen synthase kinase-3 (GSK-3).

36.-**48**. (canceled)

* * * * *